





\$39.00 per annum  
\$18.00 per volume  
Single Numbers \$9.00

BIMONTHLY  
Three volumes per year  
Immediate publication

January, 1973  
Volume 85, First Half

Founded in 1935 by Carl Murchison

# The Journal of PSYCHOLOGY

The General Field of Psychology

POWELL MURCHISON  
*Managing Editor*

HELEN M. DAVIS  
*Associate Editor*

## Editorial Board

JOHN A. ANTOINETTI  
Life Insurance Agency  
Management Association

S. HOWARD BARTLEY  
Michigan State University

ARTHUR G. BILLS  
University of Cincinnati

E. A. BOTT  
University of Toronto

LEONARD CARMICHAEL  
National Geographic Society

IRVIN L. CHILD  
Yale University

HERBERT S. CONRAD  
U. S. Office of Education

JOHN F. DASHIELL  
Florida State University

LEONARD W. DOOR  
Yale University

If this space should be unstamped, this is the regular library edition.  
But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not  
be resold for a period of five years from date of publication.

FRANK A. GELDARD  
Princeton University

ERNEST R. HILGARD  
Stanford University

JOHN E. HORROCKS  
Ohio State University

DAVID HOTHERSALL  
Ohio State University

HEINRICH KLÜVER  
University of Chicago

NORMAN L. MUNN  
University of Adelaide,  
Australia

WARREN H. TEICHNER  
New Mexico State  
University

LEE EDWARD TRAVIS  
Fuller Theological  
Seminary

JOSEPH ZURIN  
Psychiatric Institute,  
Columbia University

Copyright © 1973 by The Journal Press  
Provincetown, Massachusetts, U. S. A.

All rights reserved

Second-class postage paid at Provincetown, Mass.  
and additional mailing offices

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657



## Back Issues

The Journal Press maintains a complete inventory of back issues for all five journals by reprinting many of them each year. Early volumes frequently consist of multiple parts, and the cost of maintaining these volumes is naturally much higher than that for two-part volumes. So our old policy of a flat price per volume cannot be continued. The following rates are in effect: single numbers, \$9.00; two-part volumes, \$18.00; three-part volumes, \$27.00; four-part and over-four-part volumes, \$36.00. The only exceptions (special issues) are starred below.

	VOLUMES	DATES	PRICE PER VOLUME
<i>Journal of Genetic Psychology</i> (Pedagogical Seminary)			
	1-4	1891-April 1897	\$27.00
	5-37	July 1897-1930	36.00
	38*	1930	18.00
	39	1931	36.00
	40 and later	1932-	18.00
<i>Genetic Psychology Monographs</i>			
	1-22	1926-1940	36.00
	23 and later	1941-	18.00
<i>Journal of General Psychology</i>			
	1-3	1928-1930	36.00
	4*	1930	18.00
	5	1931	36.00
	6 and later	1932-	18.00
<i>Journal of Social Psychology</i>			
	1-10	1930-1939	36.00
	11 and later	1940-	18.00
<i>Journal of Psychology</i>			
	1 and later	1936-	18.00

A mailing charge of \$0.30 per volume is made for foreign deliveries only. Domestic mailings are prepaid. All remittances should be by check or money order payable in U.S. funds on a United States bank. Please do not send UNESCO coupons. Advance payment is required from all individuals and from foreign purchasers. Microfilm versions and special bindings are not available.

THE JOURNAL PRESS  
2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.



Notice: All copyrighted material reprinted from The Journal Press publications should be given an appropriate acknowledgment. No permission is necessary for brief quotations (less than 300-word total or less than 10 percent of length, whichever is least, for any one article). Permission is required for reprinting tables, figures, entire articles, or substantial portions thereof; and such permission will be contingent upon agreement by the original author or authors. Estate permission may be required. Since no fee is charged for a copyright license, the courtesy of a stamped, self-addressed envelope will be appreciated.

# The Journal of Psychology

Founded in 1935 by Carl Murchison

The Journal of Psychology was founded at the request of many psychologists throughout the world for the purpose of establishing a medium in which publication shall be immediate and in which other publication experiments may be carried out.

(1). The subscription price is \$39.00 per annum, \$18.00 per volume, \$9.00 per issue (three volumes per year).

(2). The format of each volume is such that each subscriber may tear down the assembled volume into its component articles.

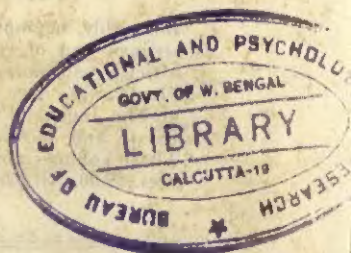
(3). Manuscripts may be sent to any member of the Editorial Board or may be sent directly to The Journal Press. The lag in publication should not be more than a few weeks.

(4). The author will receive 200 reprints of his paper, and subscribers will receive the accumulated publications bound into half-volume units. The second half of each volume will contain the title page and table of contents for that volume. Each half-volume unit will be mailed to the subscribers at its stated date.

(5). Each author will be charged the manufacturing costs of any zincs, halftone cuts, tables, or equations. Excessive use of special fonts (references, footnotes, quotations, appendices) will also be charged.

(6). The publication charges are as follows:

1 to 4 pages	\$40.00
5 to 8 "	80.00
9 to 12 "	120.00
13 to 16 "	160.00
17 to 20 "	200.00
21 to 24 "	240.00
25 to 28 "	280.00
29 to 32 "	320.00
33 to 36 "	360.00
37 to 40 "	400.00



These charges are subject to adjustment as justified.

(7). The accumulated units will be mailed to subscribers bimonthly: January, March, May, July, September, and November. The units may vary in size to accommodate publication needs.

All back issues are available. See announcement on inside front cover.

THE JOURNAL PRESS

2 Commercial Street

Provincetown, Massachusetts 02657

U. S. A.

Bureau of Ednl. & Psyl. Research  
(S. C. E. R. T.)

Date.....

Acc..... J. 8 92



# Journal of Genetic Psychology

Founded by G. Stanley Hall in 1891; reorganized by Carl Murchison in 1925

Quarterly, beginning with March. Devoted to developmental and clinical psychology, with briefly reported Replications and Refinements and occasional book reviews. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$26.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

---

# Journal of General Psychology

Founded in 1927 by E. B. Titchener and Carl Murchison

Quarterly, beginning with January. Devoted to experimental, physiological, and comparative psychology, with briefly reported Replications, Refinements, and Comments. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$26.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

---

# Journal of Social Psychology

Founded in 1929 by John Dewey and Carl Murchison

Bimonthly, beginning with February. Devoted to studies of persons in group settings and of culture and personality; special attention to cross-cultural articles and notes, and to briefly reported Replications and Refinements. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Three volumes annually. Per annum, \$39.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

---

*Each author will receive 100 unbound reprints gratis. Additional reprints may be purchased at a nominal cost when the author returns page-proof.*

THE JOURNAL PRESS

2 Commercial Street

Provincetown, Massachusetts 02657

U. S. A.

# Club Rate for Five International Psychological Journals

	PER ANNUM
THE JOURNAL OF GENETIC PSYCHOLOGY . . . . . (Quarterly, two volumes per year)	\$26.00
GENETIC PSYCHOLOGY MONOGRAPHS . . . . . (Quarterly, two volumes per year)	\$26.00
THE JOURNAL OF GENERAL PSYCHOLOGY . . . . . (Quarterly, two volumes per year)	\$26.00
THE JOURNAL OF SOCIAL PSYCHOLOGY . . . . . (Bimonthly, three volumes per year)	\$39.00
THE JOURNAL OF PSYCHOLOGY . . . . . (Bimonthly, three volumes per year)	\$39.00
	TOTAL \$156.00
TEACHERS AND STUDENTS EDITION, <i>Restricted Sale</i>	\$50.00

The *Teachers and Students Edition* is open to any individual psychologist, teacher, student, or physician but requires an agreement that the edition will not be offered for sale till after five years from publication and an agreement that the edition will not be substituted for an institutional subscription. Write for our Teachers and Students contract card.

THE JOURNAL PRESS  
2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.

## Genetic Psychology Monographs

Founded in 1925 by Carl Murchison

Quarterly, beginning with February. Devoted to developmental and clinical psychology. Each number one or more complete researches. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$26.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

*NOTICE. Effective immediately we offer prior publication for certain selected monographs. Charges will be based on the announced rates for The Journal of Psychology. Regular monographs will carry no page charge for the first 50 pages, but for all pages above 50 there will be a charge of \$10.00 per page. Manufacturing costs of any zincs, halftone cuts, tables, or equations will be charged to authors, as will excessive use of special fonts. Two hundred reprints of each monograph will be supplied free of charge.*

*The listing of monographs for 1926 through 1963 will be printed in the first half of each volume of all five journals. Monographs published after 1963 will be listed in the second half of each volume.*



# Genetic Psychology Monographs (continued)

VOLUME 1—1926

1. Performance tests for children of pre-school age—R. STUTSMAN
2. An experimental study of the eidetic type—H. KLÜVER
- 3 & 4. A study of natio-racial mental differences—N. D. M. HIRSCH
5. A psychological study of juvenile delinquency by group methods—J. W. BRIDGES AND A. M. B. BRIDGES
6. The influence of puberty precocx upon mental growth—A. GASELL

VOLUME 2—1927

- 1 & 2. The mind of a gorilla—R. M. YERKES
3. The role of eye-muscles and mouth-muscles in the expression of the emotions—K. DUNLAP
4. Family similarities in mental-test abilities—R. R. WILLOUGHBY
5. Coordination in the locomotion of infants—L. H. BURNSIDE
6. The mind of a gorilla: Part II. Mental development—R. M. YERKES

VOLUME 3—January-June, 1928

1. An experimental study of the olfactory sensitivity of the white rat—J. R. LIGGETT
2. A photographic study of eye movements in reading formuar M. A. TINKER
3. An experimental study of the East Kentucky mountaineers—N. D. M. HIRSCH
4. Responses of foetal guinea pigs prematurely delivered—G. T. AVERY
5. Objective differentiation between three groups in education (teachers, research workers, and administration)  
M. B. JENSEN
6. The effect of segregation on the sex behavior of the white rat as measured by the obstruction method—M. JENSEN

VOLUME 4—July-December, 1928

1. Observation and training of fundamental habits in young children—E. A. BOTT, W. E. BAATZ, N. CHANTY, H. BOTT
- 2 & 3. Determination of a content of the course in literature of a suitable difficulty for junior and senior high school students—M. C. BURCH
- 4 & 5. Methods for diagnosis and treatment of cases of reading disability—M. MONROE
6. The relative effectiveness of lecture and individual reading as methods of college teaching—E. B. GREENE

VOLUME 5—January-June, 1929

1. The age factor in animal learning: I. Rats in the problem box and the maze—C. P. STONE
2. The effect of delayed incentive on the hunger drive in the white rat—E. L. HAMILTON
3. Which hand in the eye of the blind?—J. M. SMITH
4. The effect of attitude on free word association-time—A. G. EKDAHL
5. The localization of tactual space: A study of average and constant errors under different types of localization  
L. E. COLE
6. The effects of gonadectomy, vasotomy, and injections of placental and orchic extracts on the sex behavior of the white rat—H. W. NISSEN

VOLUME 6—July-December, 1929

1. Learning and growth in identical infant twins: An experimental study by the method of co-twin control—GASELL AND H. THOMPSON
2. The age factor in animal learning: II. Rats on a multiple light discrimination box and a difficult maze—C. P. STONE
3. The acquisition and interference of motor habits in young children—E. MCGINNIS
4. A vocational and socio-educational survey of graduates and non-graduates of small high schools of New England—A. D. MUELLER
- 5 & 6. A study of the smiling and laughing of infants in the first year of life—R. W. WASHBURN

VOLUME 7—January-June, 1930

1. Tensions and emotional factors in reaction—E. DUFFY
2. Teacher influence on class achievement: A study of the relationship of estimated teaching ability to pupil achievement in reading and arithmetic—H. R. TAYLOR
- 3 & 4. A study of the effect of inverted retinal stimulation upon spatially coordinated behavior—P. H. EWERT
5. A study of the mental development of children with lesion in the central nervous system—E. E. LOHD
6. An experimental study upon three hundred school children over a six-year period—N. D. M. HIRSCH

VOLUME 8—July-December, 1930

1. The amount and nature of activities of newborn infants under constant external stimulating conditions during first ten days of life—O. C. IAWIN
2. Race and social differences in performance tests—S. D. PORTEUS, et al.
3. Language and growth: The relative efficacy of early and deferred vocabulary training, studied by the method of co-twin control—L. C. STRAYER
4. Eye-movements and optic nystagmus in early infancy—J. M. MCGINNIS
- 5 & 6. Reactions of kindergarten, first-, and second-grade children to constructive play materials—L. FARWELL

VOLUME 9—January-June, 1931

- 1 & 2. The status of the first-born with special reference to intelligence—H. H. HSIAO
- 3 & 4. An experimental study of bright, average, and dull children at the four-year mental level—H. P. DAVIDSON
5. An historical, critical, and experimental study of the Seashore-Kawakusa test battery—P. R. FARNSWORTH
6. A comparison of difficulty and improvement in the learning of bright and dull children in reproducing a dactyl selection—F. T. WILSON

VOLUME 10—July-December, 1931

1. A comparative study of a group of southern white and negro infants—M. B. MCGRAW
- 2 & 3. An experimental study of prehension in infants by means of systematic cinema records—H. M. HALVORSON
4. The limits of learning ability in kittens—A. M. SHURY
- 5 & 6. The effect of habit interference upon performance in maze learning—O. W. ALM

VOLUME 11—January-June, 1932

1. General factors in transfer of training in the white rat—T. A. JACKSON
2. The effect of coldy on visual apprehension and perception—M. A. TINKER
3. The reliability and validity of maze experiments with white rats—R. LEEPER
4. A critical study of two lists of best books for children—F. K. SHUTTLEWORTH
- 5 & 6. Measuring human energy cost in industry: A general guide to the literature—R. M. PAGE

VOLUME 12—July-December, 1932

1. Family resemblances in verbal and numerical abilities—H. D. CARTER
2. The development of fine prehension in infancy—B. M. CASTNER
- 3 & 4. The growth of adaptive behavior in infants: An experimental study at seven age levels—H. M. RICHARDSON
- 5 & 6. Differential reactions to taste and temperature stimuli in newborn infants—K. JENSEN

VOLUME 13—January-June, 1933

1. A critique of sublimation in males: A study of forty superior single men—W. S. TAYLOR
2. A study of the nature, measurement, and determination of hand preference—H. L. KOCH, et al.
3. The growth and decline of intelligence: A study of a homogeneous group between the ages of ten and sixty—JONES AND H. S. CONRAD
4. The relation between the complexity of the habit to be acquired and the form of the learning curve in children—M. L. MATTHEWSON
5. Eating habits in relation to personality development of two- and three-year-old children: A study of children in two nursery schools—A. A. ELIOT
6. Coordinating mechanisms of the spinal cord—O. C. INGEBRIGTSEN

# Genetic Psychology Monographs (continued)

## VOLUME 14—July-December, 1933

1. Mental growth during the first three years: A developmental study of sixty-one children by repeated tests—N. BAYLEY
2. A study of triplets: including theories of their possible genetic relationships—F. N. ANDERSON AND N. V. SCHEIDSMANN
3. The objective measurement of emotional reactions—H. V. GASKILL
4. Development of behavior in the fetal cat—J. D. CORONIOS
5. A study of certain language developments of children in grades four to twelve, inclusive—L. L. LABRANT
6. The effect of early and delayed practice on memory and motor performances studied by the method of co-twin control—J. R. HILGARD

## VOLUME 15—January-June, 1934

1. Studies in the psychology of tone and music—P. R. FARNSWORTH
2. Motor learning of children in equilibrium in relation to nutrition—E. L. BEERE
3. Discrimination limits of pattern and size in the goldfish *Carassius auratus*—J. B. ROWLEY
4. Limits of learning ability in the white rat and the guinea pig—B. F. RIESS
- 5 & 6. The limits of learning ability in rhesus monkeys—H. A. FIELD

## VOLUME 16—July-December, 1934

1. A statistical study of ratings on the California Behavior Inventory for Nursery-School Children—H. S. CONRAD
2. An eye-movement study of objective examination questions—A. FRANDSEN
3. An experimental study of constitutional types—O. KLIMBERG, S. E. ASCH, AND H. BLOCK
4. The development of a battery of objective group tests of manual laterality, with the results of their application to 1300 children—W. N. DUROST
- 5 & 6. An experimental study in the prenatal guinea-pig of the origin and development of reflexes and patterns of behavior in relation to the stimulation of specific receptor areas during the period of active fetal life—L. CARMICHAEL

## VOLUME 17—January-December, 1935

1. Organization of behavior in the albino rat—R. L. THORNDIKE
2. Brightness discrimination in the rhesus monkey—M. P. CRAWFORD
3. The limits of learning ability in cebus monkeys—A. M. KOCH
4. Nature-nurture and intelligence—A. M. LEAHY
5. On intelligence of epileptic children—E. B. SULLIVAN AND L. GARAGAN
6. A study of the play of children of preschool age by an unobtrusive observer—D. L. COCKRELL

## VOLUME 18—January-December, 1936

1. Sex differences in variational tendency—O. MCNEUMAN AND L. M. TERMAN
2. The process of learning to dress among nursery-school children—C. B. KEY, M. R. WHITE, M. P. HONZIK, A. B. HEINLEY, AND D. ERWIN
3. A study of the present social status of a group of adults, who, when they were in elementary schools, were classified as mentally deficient—W. R. BALLER
4. The influence of specific experience upon mental organizations—A. ANASTASI
- 5 & 6. Studies in aggressiveness—L. BENDER, S. KEISER, AND P. SCHILDER

## VOLUME 19—January-December, 1937

1. Psychological bases of self-mutilation—C. DABROWSKI
2. Masculine temperament and secondary sex characteristics: A study of the relationship between psychological and physical measures of masculinity—H. GILKINSON
3. A psychological study of forty unmarried mothers—R. D. NOTTINGHAM
4. Behavior problems in the children of psychotic and criminal parents—L. BENDER
5. Domination and integration in the social behavior of young children in an experimental play situation—H. H. ANDERSON
6. The sequential patterning of prone progression in the human infant—L. B. AMES

## VOLUME 20—January-December, 1938

1. The relationship between characteristics of personality and physique in adolescents—P. S. DE Q. CAROT
2. Behavior problems of elementary school children: A descriptive and comparative study—L. Y. MASTEN
3. Graphic representation of a man by four-year-old children in nine prescribed drawing situations—P. F. GRIDLEY
4. Differences between two groups of adult criminals—R. S. TOLMAN
5. A comparative study by means of the Rorschach method of personality development in twenty pairs of identical twins—E. THOMP
- Individual differences in the facial expressive behavior of preschool children: A study by the time-sampling method—C. SWAN

## VOLUME 21—January-December, 1939

1. An experimental analysis of "level of aspiration"—R. GOULD
2. Some light on the problem of bilingualism as found from a study of the progress in mastery of English among preschool children of non-American ancestry in Hawaii—M. E. SMITH
3. Domination and social integration in the behavior of kindergarten children and teachers—H. H. ANDERSON
4. The capacity of the rhesus and cebus monkey and the gibbon to acquire differential response to complex visual stimuli—W. E. GALT
5. The social-sex development of children—E. H. CAMPBELL

## VOLUME 22—January-December, 1940

1. Measuring human relations: An introduction to the study of the interaction of individuals—E. D. CHAPPLE
2. Aggressive behavior in young children and children's attitudes toward aggression—M. D. FITE
3. Student attitudes toward religion—E. NELSON
4. The prediction of the outcome-on-furlough of dementia praecox patients—J. S. JACOB
5. Significant characteristics of preschool children as located in the Conrad inventory—K. H. READ
6. Learning by children at noon-meal in a nursery school: Ten "good" eaters and ten "poor" eaters—J. B. MCCAY, E. B. WARING, AND P. J. KRAUSE

Studies in the interpretation of play: I. Clinical observation of play disruption in young children—E. H. ERIKSON

## VOLUME 23—January-June, 1941

1. An analysis of certain variables in a developmental study of language—F. M. YOUNG
2. Infant development under conditions of restricted practice and of minimum social stimulation—W. DENNIS
3. An analysis of the mental factors of various age groups from nine to sixty—B. BALINSKY
4. Factors influencing performance on group and individual tests of intelligence: I. Rate of work—M. W. BENNETT
5. Individual differences in apperceptive reaction: A study of the response of preschool children to pictures—E. W. AMEN

## VOLUME 24—July-December, 1941

1. Twins T and C from infancy to adolescence: A biogenetic study of individual differences by the method of co-twin control—A. GESKILL AND H. THOMPSON
2. Finger nail-biting: Its incipency, incidence, and amelioration—A. L. BILLIG
3. An experimental study of the factors of maturation and practice in the behavioral development of the embryo of the frog, *Rana pipiens*—A. FROMME
4. The Fels child behavior scales—T. W. RICHARDS AND M. P. SIMONS
5. Measurement of the size of general English vocabulary through the elementary grades and high school—M. K. SMITH
6. Stereotypes in the field of musical eminence—P. R. FARNSWORTH

## VOLUME 25—January-June, 1942

1. A study of factors determining family size in a selected professional group—J. C. FLANAGAN
2. A genetic study of geometrical-optical illusions—A. WALTERS
3. Interpretation of behavior-ratings in terms of favorable and unfavorable deviations: A study of scores from the Read-Conrad Behavior Inventory—K. H. READ AND H. S. CONRAD
4. Are there any innate behavior tendencies?—J. B. SCHOELLAND
5. An investigation of the intelligibility of the speech of the deaf—C. V. HODGINS AND F. C. NUMBERS

# Genetic Psychology Monographs (continued)

VOLUME 26—July-December, 1942

1. The critical frequency limen for visual flicker in children between the ages of 6 and 18—V. L. MULLER
2. Some factors determining handedness in the white rat—K. L. WENTWORTH

2. Motivation and behavior—E. FERNER-BRUNSWICK

VOLUME 27—January-June, 1943

1. Comparison of children's personality traits, attitudes, and intelligence with parental occupation—N. R. MADHY
2. A comparative study of mental functioning patterns of problem and non-problem children seven, eight, and nine years of age—M. L. PIGNATELLI

VOLUME 28—July-December, 1943

1. Separation anxiety in young children: A study of hospital cases—H. EDLSTON
2. Correlates of vocational preferences—W. A. BRADLEY, JR.

VOLUME 29—January-June, 1944

1. Mental changes after bilateral prefrontal lobotomy—S. D. PORTER and R. D. KEPNER
2. A twin-controlled experiment on the learning of auxiliary languages—B. PRICE, W. J. KOSTER, and W. M. TAYLOR

VOLUME 30—July-December, 1944

1. A method of administering and evaluating the thematic apperception test in group situations—R. M. CLARK
2. A study of anxiety reactions in young children by means of a projective technique—R. TEMPLE and E. W. AMEN

VOLUME 31—January-June, 1945

1. The evolution of intelligent behavior in rhesus monkeys—B. WEINSTEIN
2. Perceptual behavior of brain-injured, mentally defective children: An experimental study by means of the Rorschach technique—H. WERNER

VOLUME 32—July-December, 1945

1. A clinical study of sentiments: I.—H. A. MURRAY and C. D. MORGAN
2. A clinical study of sentiments: II.—H. A. MURRAY and C. D. MORGAN

VOLUME 33—January-June, 1946

1. Interpretation of spontaneous drawings and paintings—T. S. WERNER
2. Preferences for sex symbols and their personality correlates—K. FRANGE
2. Outstanding traits: In a selected college group, with some reference to career interests and war records—F. L. TAYLOR and W. L. WOODS

VOLUME 34—July-December, 1946

1. The relation of emotional adjustment to intellectual function—J. L. DESPLET and H. O. PIERCE
2. The smiling response: A contribution to the ontogenesis of social relations—R. A. SPITZ
2. Finger-painting and personality diagnosis—P. J. NAPOLI

VOLUME 35—January-June, 1947

1. The thematic apperception technique in the study of culture-personality relations—W. E. HENRY
2. A continuation study of anxiety reactions in young children by means of a projective technique—M. DORNEY and E. W. AMEN
- A study of the vocational interest trends of secondary school and college women—A. M. CAWLEY

VOLUME 36—July-December, 1947

1. Maze test validation and psychosurgery—S. D. PORTER and H. N. PERKINS
2. The diagnostic implications of Rorschach's test in case studies of mental defectives—I. JOLLES

VOLUME 37—January-June, 1948

1. The radio day time serial: A symbolic analysis—W. L. WARNER and W. E. HENRY
2. The relation of personality characteristics and response to verbal approval in a learning task—G. I. GRACE
2. The mechanism of vision: XVIII. Effects of destroying the visual "associative areas" of the monkey—K. S. LASHLEY
- A study of the relationship between handwriting and personality variables—P. CASTELLINO and T. HEROUX

VOLUME 38—July-December, 1948

1. Modern language learning: The intensive course as sponsored by the United States Army, and implications for the undergraduate course of study—M. LIND
2. Conflict: A study of some interactions between appetite and aversion in the white rat—M. A. TOLCOTT
2. Schizophrenia and the MAPS test: A study of certain formal psycho-social aspects of fantasy production in schizophrenia as revealed by performance on the Make a Picture Story (MAPS) Test—F. S. SHKIDMAN
- A study of the transmission of authority patterns in the family—H. L. INGERSOLL

VOLUME 39—January-June, 1949

1. A study of the psychoanalytic theory of psychosexual development—C. S. BLUM
2. The assessment of parental attitudes in relation to child adjustment—E. J. SHOREN, JR.
2. Qualitative differences in the vocabulary responses of normals and abnormals—H. FRIZEL
- The relative effectiveness of motion and still pictures as stimuli for eliciting fantasy stories about adolescent-parent relationships—P. E. EISENBERG
- The organization of hereditary maze-brightness and maze-dullness—L. V. SEARLE

VOLUME 40—July-December, 1949

1. An experimental study of what young school children expect from their teachers—B. BIERER and C. LEWIS
2. A study of the relative effects of age and of test difficulty upon factor patterns—H. A. CURTIS
2. A projective experiment using incomplete stories with multiple choice endings—J. K. SEATON
2. Effects of sex role and social status on the early adolescent personality—E. MILNER
- Social perceptions and attitudes of children—M. RADKE, H. TRAGER, and H. DAVIS

VOLUME 41—January-June, 1950

1. Some psychological and educational aspects of pediatric practice: A study of well-baby clinics—L. H. BLUM
2. One-trial learning in the domestic rat—B. B. HUDSON
2. An introduction to the principles of scientific psychoanalysis—A. ELLIS
2. Awareness of racial differences by preschool children in Hawaii—D. V. SPRINGER
2. Age trends in children's evaluation of teacher-approved and teacher-disapproved behavior—S. L. WITKOV
- The relationship between level of vocational aspiration and certain personal data: A study of some traits and influences bearing on the prestige level of vocational choice—J. STEUBINS

VOLUME 42—July-December, 1950

1. Personality patterns of suicidal mental hospital patients—N. L. FARBEROW
2. Sex-role identification in young children in two diverse social groups—M. RABANAN
2. A study of the influence of the social field on individual behavior: As revealed in the expression of hostility and warmth by neurotics and paranoid schizophrenics in discussion group situations—D. SHAPIRO
- An experimental study of avoidance—R. F. HEFFERLINE

VOLUME 43—January-June, 1951

1. A study of copying ability in children—E. A. TOWNSEND
2. Prestige motivation of gifted children—D. P. AUSUBEL
2. A psychological study of physical scientists—A. ROX

VOLUME 44—July-December, 1951

1. The organization of hostility controls in various personality structures—S. FISHER and E. HINDS
2. Children and radio: A study of listeners and non-listeners to various types of radio programs in terms of selected ability, attitude, and behavior measures—E. A. RICCIUTI
2. Quantitative expression in young children—W. E. MARTIN
- The use of magnetic devices in the collection and analysis of the preverbal utterances of an infant—A. W. LYNN

VOLUME 45—January-June, 1952

1. Japanese-American personality and acculturation—W. CAUDILL
2. A statistical study of the Freudian theory of levels of psychosexual development—C. A. BARNES
2. Personality characteristics of selected disability groups—D. N. WIENER



# Genetic Psychology Monographs (continued)

VOLUME 46—July-December, 1952

1. The relationship of social status, intelligence, and sex of ten- and eleven-year-old children to an awareness of poverty—F. J. ESTYAN  
An empirical study of the castration and Oedipus complexes—S. M. FRIEDMAN
2. The relationship between projective test scoring categories and activity preferences—M. M. SCHWARTZ  
A comparison of formal and content factors in the diagnostic testing of schizophrenia—M. SHERMAN

VOLUME 47—January-June, 1953

1. Ability and accomplishment of persons earlier judged mentally deficient—D. C. CHARLES  
Variations in the consistency of the behavioral meaning of personality test scores—M. KORNREICH
2. Some child-rearing antecedents of aggression and dependency in young children—R. R. SEARS, *et al.*  
Symptom correlates for descriptive diagnosis—J. R. WITTENBORN, *et al.*

VOLUME 48—July-December, 1953

1. Age and mental abilities: A longitudinal study—W. A. OWENS, JR.  
The development of a personality questionnaire for drinkers—P. J. HAMPTON  
Personality and physical disease: A test of the Dunbar hypothesis applied to diabetes mellitus and rheumatic fever—D. H. CROWELL
2. Socio-economic contrasts in children's peer culture prestige values—B. POPE  
A critical review of the stability of social acceptability scores obtained with the partial-rank-order and the paired-comparison scales—S. A. WITKOL AND G. C. THOMPSON  
A study of the effects of color on Rorschach responses—G. C. BROOY

VOLUME 49—January-June, 1954

1. Factors underlying major reading disabilities at the college level—J. A. HOLMES  
Parent behavior toward first and second children—J. K. LASKO
2. Social-status and intelligence: An experimental study of certain cultural determinants of measured intelligence—E. A. HAGGARD  
Certain determinants and correlates of authoritarianism—S. SECKL  
Personalities in faces: I. An experiment in social perceiving—P. F. SECORD, W. F. DUKES, AND W. BRYAN

VOLUME 50—July-December, 1954

1. A study of the relationship between play patterns and anxiety in young children—E. W. AMEN AND N. RKNISON  
Operational exploration of the conceptual self system and of the interaction between frames of reference—M. EDELSON AND A. E. JONES  
Problem solving: A statistical description of some relationships between organismic factors and selected response measures—N. A. FATTU, E. KAPOS, AND E. V. MECH
2. The relation of cortical potentials to perceptual functions—C. CHVATTE  
The import for clinical psychology of the use of tests derived from theories about infantile sexuality and adult character—D. W. MILES  
Measuring personality in developmental terms: The Personal Preference Scale—M. H. KROUT AND J. K. TABIN

VOLUME 51—January-June, 1955

1. Some relations between techniques of feeding and training during infancy and certain behavior in childhood—A. BERNSTEIN  
The expression of personality in drawings and paintings—L. H. STEWART
2. Negative stereotypes concerning Americans among American-born children receiving various types of minority-group education—J. A. FISHMAN  
The Lincoln-Oseretsky Motor Development Scale—W. SLOAN

VOLUME 52—July-December, 1955

1. Some personality correlates of sex, sibling position, and sex of sibling among five- and six-year-old children—H. L. KOCH  
A quantitative Rorschach assessment of maladjustment and rigidity in acculturating Japanese Americans—G. DeVOS  
Measurement of authoritarianism and its relation to teachers' classroom behavior—H. M. MCGEE
2. The formal aspects of schizophrenic verbal communication—B. MINTZ  
A study in an aspect of concept formation, with subnormal, average, and superior adolescents—H. N. HOFFMAN  
Traumatic avoidance learning: Acquisition and extinction in dogs deprived of normal peripheral autonomic function—L. C. WYNN AND R. L. SOLOMON

VOLUME 53—January-June, 1956

1. As the psychiatric aide sees his work and problems—F. L. WELLS, M. GREENBLATT, AND R. W. HYDE  
An investigation of avoidance, anxiety, and escape behavior in human subjects as measured by action potentials in muscle—J. D. BROTHERS  
Spread of effect: A critical review—M. H. MARK
2. Stress, fantasy, and schizophrenia: A study of the adaptive processes—O. J. B. KERNER  
The attitude structure of the individual: A Q-study of the educational attitudes of professors and laymen—F. N. KERLINGER

VOLUME 54—July-December, 1956

1. A study of personality differences between middle and lower class adolescents: The Second Test in culture-personality research—L. RAINWATER  
The assessment of parental identification—S. W. GRAY AND R. KLAUS
2. The influence of social context on impulse and control tendencies in preadolescents—G. H. ZUK  
Tender-mindedness versus tough-mindedness in psychology: A reexamination—H. WINTHROP  
A method for the comparison of groups: A study in thematic apperception—L. C. SCHRAW AND W. E. HENRY

VOLUME 55—January-June, 1957

1. Academic performance and personality adjustments of highly intelligent college students—B. M. HORRALL  
The use of the Vineland Social Maturity Scale in the planning of an educational program for non-institutionalized low-grade mentally deficient children—M. G. GORTSCHEK
2. The structure and origin of the anal character—H. BELOFF  
Free expression of adolescents' interests—M. AMATORA  
The role of mass media and the effect of aggressive film content upon children's aggressive responses and identification choices—R. S. ALBERT  
Interest in persons as an aspect of sex difference in the early years—E. W. GOODENOUGH

VOLUME 56—July-December, 1957

1. Some stable response determinants of perception, thinking, and learning: A study based on the analysis of a single test—M. L. SIMMEL AND S. COUNTS  
Potentials of age: An exploratory field study—S. L. PRESSAT
2. Tactile communication—L. K. FRANK  
The use of a filmed puppet show as a group projective technique for children—M. R. HAWORTH  
The social competence of middle-aged people—R. J. HAVIGHURST

VOLUME 57—January-June, 1958

1. Psychological and cultural problems in mental subnormality: A review of research—S. B. SARASON AND T. GLADWIN
2. Developmental aspects of discrimination in relation to adjustment—P. LYNDON  
Muscular tension as a response to psychological stress in rheumatoid arthritis and peptic ulcer—J. A. SMITHWORTH  
Frustration: Theory and experiment—R. LAWSON AND M. H. MARK

VOLUME 58—July-December, 1958

1. Emotional aspects of political behavior: The women voter—E. M. BENNETT AND H. M. GOODWIN  
The accuracy of self estimate: A measure of self-concept reality—R. M. BRANDT  
Personality factors in social mobility: A study of occupationally mobile businessmen—J. C. AREGLEEN
2. Interaction between authoritarian and nonauthoritarian principals and teachers—P. LAMBERT  
Some factors related to the choice-status of ninety eighth-grade children in a school society—D. ELKINS  
Some relationships between fantasy and self-report behavior—R. C. CALOGERAS

# Genetic Psychology Monographs (continued)

## VOLUME 59—January-June, 1959

1. Freshman Rorschachs and college performance—G. R. SCHMEIDLER, M. J. NELSON, AND M. BRISTOL  
Perceptual aspects of attributed movement—R. W. COAN  
Men and women: Personality patterns and contrasts—E. M. BENNETT AND L. R. COHEN
2. Personality factors in mothers of cerebral palsied children—G. BOLES  
Separatism and integrationism: A social-psychological analysis of editorial content in New York newspapers of three American minority groups—J. A. FISHMAN AND G. S. FISHMAN  
Self, role, and satisfaction—A. L. BROPHY

## VOLUME 60—July-December, 1959

1. The constancy of personality ratings over two decades—R. D. TUDDENHAM  
Mother-son identification and vocational interest—L. H. STEWART  
A study of mother-child relationships in the emotional disorders of children—M. J. ROSENTHAL, M. FINKELSTEIN, E. NI, AND R. E. ROBERTSON  
On the trail of the wolf-children—W. F. OGBURN AND N. K. BOSE
2. Measuring the mental health of normal adults—R. F. PECK  
Parental identification in young children—W. EMMERICH  
The problems of Jewish religious youth—E. M. LEVINSON

## VOLUME 61—January-June, 1960

1. Identifying socially maladjusted school children—A. A. ADAMS  
Social interaction in an interracial nursery school—H. W. STEVENSON AND N. G. STEVENSON, *et al.*  
Interests of pre-adolescent boys and girls—M. AMATOGA  
Educational and sex differences in the organization of abilities in technical and academic students in Colombia, South America—J. F. FILELLA
2. Some themes in the personalities of German men—L. RAINWATER  
Right-left gradients in body image, body reactivity, and perception—S. FISHER  
Longitudinal survey of child Rorschach responses: Younger subjects two to 10 years—L. B. AMES  
Personality patterns related to occupational roles—M. SIEGELMAN AND R. F. PECK

## VOLUME 62—July-December, 1960

1. The reading process: A descriptive, interdisciplinary approach—M. P. LANGMAN  
Reactions to stress as a function of the level of intelligence—S. M. H. ZAIDI  
Rorschach diagnosis by a systematic combining of content, thought process, and determinant scales—P. A. BOWER, R. TESTIN, AND A. ROBERTS  
Longitudinal survey of child Rorschach responses: Older subjects aged 10 to 16 years—L. B. AMES
2. Twenty years of shock therapy in America, 1937-1956: An annotated bibliography—N. H. PRONKO, R. SITTERLY, AND K. BEAC  
The quantitative analysis of parent behavior toward psychotic children and their siblings—E. M. DONNELLY

## VOLUME 63—January-June, 1961

1. Make a sentence test: An approach to objective scoring of sentence completions—E. F. BORCATT  
Some aspects of the relationship between perception and motility in children—E. P. ROTHMAN  
Some psychophysical aspects of mental illness in children—P. H. BERKOWITZ  
Conceptual thinking in schizophrenic children—G. FRIEDMAN
2. Some maternal influences on children's personality and character—J. C. FINNEY  
A study of concept formation in normals, mental defectives, and brain-damaged adults—R. W. ZASLOW  
An investigation of trait relationships among six-year-old children—S. A. AVAKIAN

## VOLUME 64—July-December, 1961

1. Coping patterns of preschool children in response to intelligence test demands—A. MORIARTY  
Studies of individual and paired interactional problem-solving behavior of rats: II. Solitary and social controls—A. BARON AND R. A. LITTMAN
2. Studies of role performance—J. H. MANN  
A psychological assessment of professional actors and related professions—R. TAFT  
A cross-cultural study of menstrual taboos—W. N. STEPHENS  
The biological roots of creativity—H. GUTMAN

## VOLUME 65—January-June, 1962

1. A Rorschach study of the development of personality structure in white and Negro children in a Southeastern community—A. C. PRICE  
Measuring personality patterns of women—J. LOKYNGER  
Home interests in early adolescence—M. AMATOGA  
Social and emotional adjustment during adolescence as related to the development of psychosomatic illness in adulthood—L. H. STEWART
2. A systematic study of interaction process scores, peer and self-assessments, personality and other variables—E. F. BORCATT  
Children's conceptions of the content and functions of the human body—E. GELLERT

## VOLUME 66—July-December, 1962

1. Ocular-manual laterality and perceptual rotation of literal symbols—J. FLEISCHER  
Parental child-rearing attitudes and their correlates in adolescent hostility—S. B. CHOROST  
Some explorations into the nature of anxieties relating to illness—P. R. ROBBINS  
Developmental-interactive aspects of hypnosis: Some illustrative cases—J. R. HILGARD AND E. R. HILGARD
2. Teaching a two-year-old to read: An experiment in early childhood learning—W. FOWLER  
Achievement gains or losses during the academic year and over the summer vacation period: A study of trends in achievement by sex and grade level among students of average intelligence—K. M. PARSLEY, JR., AND M. POWELL

## VOLUME 67—January-June, 1963

1. Verbatim dialogue of a mother and child in therapy—C. E. MOUTAKAS  
A study of self-concept: Delinquent boys' accuracy in selected self-evaluations—W. E. AMOS  
An integrating study of the factor structure of adult attitude-interests—R. B. CATTELL AND J. HORN  
Social factors in mental disorders: A contribution to the etiology of mental illness—W. C. COBLINER
2. Conceptual preferences associated with choice of quantitative and nonquantitative college majors—J. GILBERT  
A study of the concept of the feminine role of 51 middle-class American families—A. STEINMANN

## VOLUME 68—July-December, 1963

1. Brain dynamism as reflected in illusions—G. K. YACOVYNSKI  
The nature and measurement of components of motivation—R. B. CATTELL, J. A. RADCLIFFE, AND A. B. SWENEY
2. Sharing in preschool children as a function of amount and type of reinforcement—W. F. FISCHER  
The Gesell Incomplete Man Test as a measure of developmental status—L. B. AMES AND F. L. ILC  
Psychological and sociological factors in prediction of career patterns of women—M. C. MULVEY  
Functionalism and its historical significance—R. HARRISON





## PUBLICATION IN THE JOURNAL OF PSYCHOLOGY

*The Journal of Psychology* provides a special service to authors and readers. Articles accepted for publication appear in print with the least possible delay. Although editorial standards are the same as those for the other journals, all correspondence and handling are expedited, and the referees may, at their discretion, permit some latitude in length of treatment with respect to tabular material, discussion, and review of the literature.

To save time, authors are requested to exercise unusual care in preparing their papers. Special attention should be paid to the General Instructions on the back cover of each issue and to the announcement on the first preliminary page.

### *Special Announcement*

Several changes, some recent and some effective with 1973 issues, have been made in our journals. We are now printing summaries at the beginning instead of the end of each article. The Caslon type face that we have used for many years is no longer available, and we have substituted Old Style No. 7, which prints exactly the same number of letters per line but gives increased readability. Our new text paper reduces the bulk of each issue by about 20 percent while increasing opacity, and the perfect binding has many advantages over the old side wiring, such as elimination of rusting with age, saving of bulk, and ease of opening flat. We hope that everybody likes the modernization.



\$39.00 per annum  
\$18.00 per volume  
Single Numbers \$9.00

BIMONTHLY  
Three volumes per year  
Immediate publication

January, 1973  
Volume 83, First Half

(Founded in 1935 by Carl Murchison)

# The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a bilateral contract, and may not be resold for a period of five years from the date of publication.

JANUARY, 1973

(Manuscripts are printed in the order of final acceptance)

Developmental trends in personal space . . . . .	3
BY DARHL M. PEDERSEN	
A comparison of two kinds of dogmatism scores: Rokeach categories <i>versus</i> open-ended responses . . . . .	11
BY MARION STEININGER	
A comparison of husband-wife responses to pregnancy . . . . .	17
BY STEWART MEIKLE AND RICHARD GERRITSE	
Creativity: Performance, profiles, and perceptions . . . . .	25
BY VICTOR K. PHILLIPS	
Awareness in verbal operant conditioning: Examination of performance quality changes . . . . .	31
BY EVERETT E. ADAM, JR., AND ROBERT J. PAUL	

(OVER)

Copyright © 1973 by The Journal Press  
Provincetown, Massachusetts, U. S. A.

All rights reserved

Second-class postage paid at Provincetown, Mass  
and additional mailing offices

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

Cue utilization patterns in student-faculty evaluation . . . . .	41
BY STEVEN E. PERMUT	
A child-adult research form of the Pittsburgh Scales of Social Extraversion-Introversion and Emotionality . . . . .	49
BY NICHOLAS A. SIEVEKING	
Behavioral variability among retardates, children, and college students . . . . .	57
BY ANN M. LIEBERT AND ALFRED A. BAUMEISTER	
Impaired sex-role identification in schizophrenia expressed in the comprehension of humor stimuli . . . . .	67
BY JONATHAN ECKER, JACOB LEVINE, AND EDWARD ZIGLER	
Relations among sensation seeking and simulated and behavioral personal space . . . . .	79
BY DARHL M. PEDERSEN	
Some correlates of the level of constraint in a system of social attitudes . . . . .	89
BY WILLIAM W. RAMBO, WARREN H. JONES, AND PHILLIP D. FINNEY	
Reactions to aggression-related stimuli following reinforcement of aggression . . . . .	95
BY RUSSELL G. GEEN AND DAVID STONNER	
So-called "nervous habits" . . . . .	103
BY D. G. WILLIAMS	
Creativity in rural, urban, and Indian children . . . . .	111
BY JOHN D. WILLIAMS, JOHANNA TEUBNER, AND STEVEN D. HARLOW	
Brainstorming in large groups as a facilitator of children's creative responses . . . . .	117
BY LYLE J. BUCHANAN, JR., AND HENRY CLAY LINDGREN	
Drug usage, personality, attitudinal, and behavioral correlates of driving behavior . . . . .	123
BY KAY JAMISON AND WILLIAM H. MCGLOTHLIN	
Congruence of adolescents' self-concepts and parents' perceptions of adolescents' self-concepts . . . . .	131
BY JOSEPH C. BLEDSOE AND R. GENE WIGGINS	
A study of the self-esteem and alienation of male homosexuals . . . . .	137
BY JERROLD S. GREENBERG	
Age, sex, and title of therapist as determinants of patients' preferences . . . . .	145
BY WILLIAM E. SIMON	
Procedures for estimating magnitude of effects . . . . .	151
BY JOHN GAITO AND JOHN FIRTH	
The dyslexic child—Two years later . . . . .	163
BY EDNA J. HUNTER AND HADLEY M. LEWIS	



## DEVELOPMENTAL TRENDS IN PERSONAL SPACE\* <sup>1</sup>

*Department of Psychology, Brigham Young University*

---

DARHL M. PEDERSEN<sup>2</sup>

---

### SUMMARY

Personal space scores were obtained for 11 males and females in each of six elementary school grades toward four stimulus persons: man, woman, boy, girl. An analysis of variance with sex, grade, and stimulus person as the main effects produced significant sex, grade, sex  $\times$  grade, and sex  $\times$  stimulus person effects. Males had a larger personal space than females at all grades and for all stimulus persons. The mean for male third graders was significantly different from all other means for males and females. A larger personal space for males than for females established by the third grade tended to remain until the sixth grade, although personal space tended to decline for both sexes. A consistent downward trend in personal space for females reversed itself in the sixth grade when the personal space for females increased to approximately the same distance as for male sixth graders.

The personal space of males and females toward men, women, and same sexed children tended to be at about the same level. However, personal space toward opposite sexed children tended to be smaller.

---

### A. INTRODUCTION

Personal space is the area surrounding a person which he regards as his own and which he does not normally like the other people to penetrate (1, 4, 11, 14). Some attempts to measure personal space have involved finding distances between *S* and other persons (2, 5, 6, 13) or finding the distances between figures representing *S* and figures representing other persons. The former measures may be called behavioral personal space measures, and the latter simulated personal space measures.

Simulated personal space measures are related to the concept of social

---

\* Received in the Editorial Office on September 22, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This experiment was supported in part by a Faculty Research Fellowship granted by Brigham Young University.

<sup>2</sup> Requests for reprints should be sent to 1230-A SFLC, Brigham Young University; Provo, Utah 84601.

schemata. Originally Kuethe (7, 8, 9) and Kuethe and Stricker (10) defined social schemata as the tendency of human Ss to group human stimulus objects together. The human stimulus objects used were felt figures. After some initial attention toward the configuration of groupings, distances between figures were investigated (9). Kuethe (9, p. 248) has indicated that social schemata are "learned during many year of social experiences." Developmental studies designed to trace the process by which social interaction distances are learned are missing from the literature. The purpose of the present study was to investigate developmental trends in personal space among elementary school children with use of a simulated measure.

## B. METHOD

### 1. Subjects

Ss were 132 students enrolled in elementary school in Milpitas, California. Eleven male and 11 female Ss were selected at random from each of the six grades from one to six.

### 2. Measuring Instrument

A new measure entitled the Pedersen Personal Space Measure—Children's Form (PPSM) was devised and used in this study because of the unavailability of a suitable stimulated personal space measure using figures of children. It was thought that children would be able to identify with figures representing a same sexed child better than with figures representing adults. There were two versions of the PPSM: one for males and one for females. Both consisted of  $24\frac{1}{2}'' \times 11''$  pages placed horizontally. On the left side of each page a standing profile of a person was printed approximately 2'' from the left and 3'' from the bottom. A line on which the profile was standing extended  $7\frac{1}{4}''$  to the right.

A moveable profile was attached with a paper clip to the upper right hand corner of each page. The profiles represented a man (M), a woman (W), a boy (B), or a girl (G). Each of these was presented facing right (R), front (F), or left (L). Ss were instructed that for the first 12 items they were "to pretend that the one that is loose is you, and that the one that is not is ~~some~~ else. It might be a man, a lady, a boy your age, or a girl your age." Ss were directed to position the moveable profile on the line so that the distance between the profile representing the self and a profile representing another person was as close as possible so that Ss still felt comfortable. Ss were told that their task was the same for the second set of 12 items except that the



self would be the stationary left profile, and another person would be the moveable right profile. For all items the moveable profile was facing left toward the stationary profile. The profiles used are presented in Figure 1.

For the first 12 items the left profiles were as follows: M-R, W-R, B-R, G-R, M-F, W-F, B-F, G-F, M-L, W-L, B-L, and G-L. The right profiles were a boy (or a girl) facing left. For the second 12 items the left profiles were in three sets. The first set had a boy (or girl) facing right. The second set had a boy (or girl) facing front. And the third set had a boy (or girl) facing left. The right profiles consisted of three sets of profiles in the order M-L, W-L, B-L, and G-L.

### 3. Procedure

The PPSM was administered individually during a 10 minute testing session. The distance for each item was measured to the nearest millimeter. A

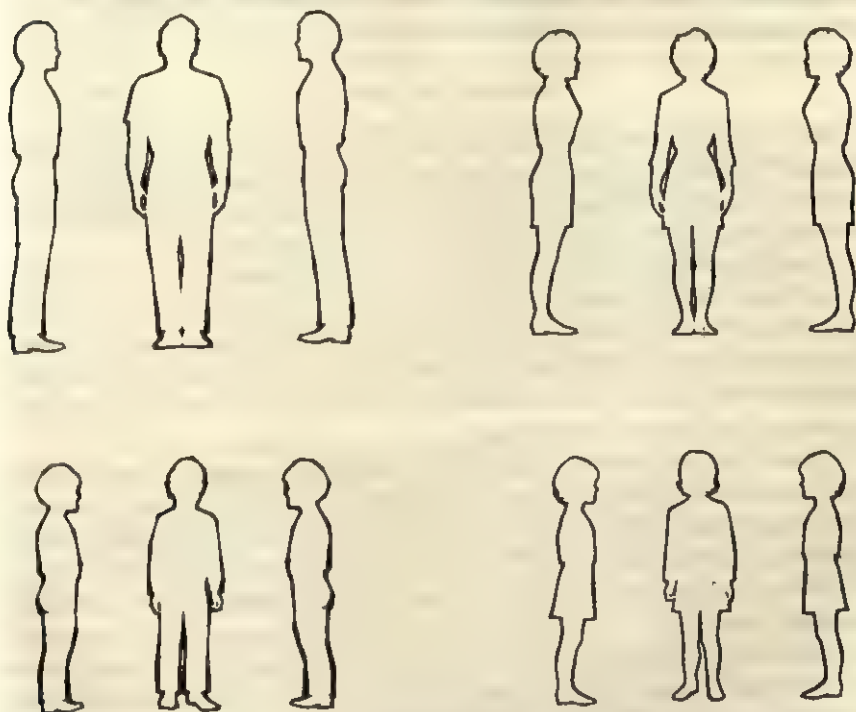


FIGURE 1  
PROFILES USED IN THE PEDERSEN PERSONAL SPACE MEASURE—CHILDREN'S FORM

preliminary factor analysis of the PPSM disclosed that the scores on the 24 items could be represented by four subscores, each representing personal space relative to a single stimulus person. The subscores were found by finding the mean distance for the six items involving the stimulus person. The four subscores were (a) man, (b) woman, (c) boy, and (d) girl.

#### 4. Data Analysis

A  $2 \times 6 \times 4$  factorial analysis of variance was completed. The main factors were sex, grade, and stimulus person. Stimulus person was a repeated measure. This analysis permitted a determination of (a) differences in personal space between boys and girls, (b) differences in personal space from one grade level to another, (c) differences in personal space from one stimulus person to another, and (d) differences in personal space produced by interactions of the main factors. A Duncan's New Multiple Range Test (15) was applied to significant mean and interaction effects to determine those means that differed significantly from the remaining means.

#### C. RESULTS AND DISCUSSION

Sex of subject was significant at the .01 level ( $F = 7.89$ ). The mean distance for females (20.47) was significantly smaller than for males (23.18). Across all grade levels and stimulus persons, females placed the profile representing them significantly closer to other figures than males did.

A significant effect at the .01 level was found for grade ( $F = 4.27$ ). The grade means were as follows: first, 20.98; second, 22.69; third, 26.16; fourth, 21.02; fifth, 18.90; sixth, 21.23. The Duncan Test revealed that the mean of third graders was significantly different from each of the others. Also the mean for fifth graders was significantly different from that of the second graders at the .05 level of significance.

A more complete understanding of the differences among the grade means is obtained by considering the two sexes separately. This is relevant, since the sex  $\times$  grade interaction was significant at the .01 level ( $F = 3.86$ ). Table 1 gives the mean distance on the PPSM for males and females by grade level.

The table shows that the larger mean for third graders resulted primarily from the large mean for male third graders. The Duncan Test showed that the mean for male third graders was significantly different from all other means for males and females at the .01 level of significance. Also, the mean for female fifth graders was significantly different from the means of male second graders (.01 level), third graders (.01 level), and fourth graders (.05). It was also

TABLE 1  
MEAN DISTANCES ON THE PEDERSEN PERSONAL SPACE MEASURE—CHILDREN'S  
FORM FOR MALES AND FEMALES FOR EACH GRADE

Grade	Males	Females
1	18.99	22.97
2	24.13	21.25
3	31.05	21.27
4	22.97	19.07
5	20.95	16.84
6	21.02	21.44

significantly different from the mean of female first graders (.05 level). Thus, the smaller mean for fifth graders mainly resulted from the low mean for female fifth graders.

An examination of the trends in the magnitude of the means showed that the male and female means were roughly parallel from the second through the fifth grades, with the male means being larger. However, there were notable reversals for the first and sixth grades, with females having larger means than males for those two grades (although the differences were not statistically significant).

A number of investigators have found that the adult female has a smaller personal space than the adult male (1, 3, 5, 12). The results of the present study indicated that this difference emerged in the elementary school years. Initially, there was no difference in personal space between males and females. In fact, females tended to have larger personal spaces than males for the first grade, although the difference was not significant. However, by the third grade males were significantly larger in personal space than females. The mean differences persisted until the fifth grade. The social learning factors that produced the larger personal space of males toward other people by the third grade are unknown.

It is clear that the larger personal space was not simply a function of the responses of the males toward female peers. The sex  $\times$  grade  $\times$  stimulus person interaction was not significant. The male third grader responded with larger personal space toward all stimulus persons. Conversely, by the fifth grade the female had developed a close personal space with all types of stimulus persons. The sharp increase in the mean personal space distance of sixth grade females to levels that existed before the third and fourth grades is inexplicable. There was a clear reversal in the downward trend in means that existed from the first through the fifth grades. An understanding of this result must await further investigation of the dynamics of the social milieu of sixth grade girls.



A significant sex  $\times$  grade  $\times$  stimulus person effect would have indicated that some of the developmental trends in personal space were ascribable to differential responses to the various stimulus persons. It was not significant. However, the sex  $\times$  stimulus person interaction was significant at the .01 level ( $F = 6.97$ ). It indicated that males and females responded differentially to the stimulus person in ways that persisted over all grades. The means of males and females for the four stimulus persons are given in Table 2. The Duncan Test showed that the personal space of females toward boys was

TABLE 2  
MEAN DISTANCES FOR EACH STIMULUS PERSON ON THE PEDERSEN PERSONAL SPACE  
MEASURE—CHILDREN'S FORM FOR MALES AND FEMALES

Stimulus person	Males	Females
Man (M)	26.19	20.59
Woman (W)	23.40	22.45
Boy (B)	24.39	16.32
Girl (G)	18.76	22.54

significantly *closer* than toward any of the other stimulus persons (.01 level for G and W, and .05 level for M). It was also significantly (.01 level) closer than the personal space of boys toward all stimulus persons except girls. Furthermore, the personal space of boys toward girls was significantly closer than toward any other stimulus persons (.01 level for M, and .05 level for B and W). This result is quite surprising, since it is commonly believed that children in this age bracket prefer to "keep their distance" from opposite sexed children of the same age. In fact, their personal space requirements were less toward opposite sexed peers than toward men, women, and same sexed peers. Responses to men, women, and same sexed peers were not significantly different for either males or females.

The last mean difference that was significant (.05 level) was between the means of males and females toward M. Females established closer distances toward a man than males did. The M-figure represented men in general. It is unknown what the results would have been if the M and W figures had represented father and mother.

The significant sex  $\times$  stimulus person interaction was produced by the small mean distances of males toward G, and females toward B. If the means were to be rearranged for females in Table 2 so that the stimulus persons were same sexed adult, opposite sexed adult, same sexed child, and opposite sexed child, respectively, the mean distances for males and females would be virtually parallel. (If this scheme had been used in the original analysis, the

significant sex  $\times$  stimulus person interaction would have disappeared, and a significant stimulus person main effect would have occurred.) Males tended to respond with greater distance toward all stimulus persons than did females. Except for the response toward opposite sexed child, responses toward the other stimulus persons tended to be at the same level for both males and females.

## REFERENCES

1. DOSEY, M. A., & MEISELS, M. Personal space and self-protection. *J. Personal. & Soc. Psychol.*, 1969, **11**, 93-97.
2. FELIPE, N., & SOMMER, R. Invasions of personal space. *Soc. Prob.*, 1966, **14**, 206-214.
3. HOROWITZ, M. F. Human spatial behavior. *Amer. J. Psychother.*, 1965, **19**, 20-28.
4. ———. Body image. *Arch. Gen. Psychiat.*, 1966, **14**, 456-460.
5. HOROWITZ, M. F., DUFF, D. F., & STRATTON, L. O. Body-buffer zone: Exploration of personal space. *Arch. Gen. Psychiat.*, 1964, **11**, 651-656.
6. KINZEL, A. Abnormalities of body buffer zone in violent prisoners. Presented at the Third Joint Meeting of the Clinical Society and Commissioned Officers Association of the U.S. Public Health Service, San Francisco, California, 1968.
7. KUETHE, J. L. Social schemas. *J. Abn. & Soc. Psychol.*, 1962, **64**, 31-38.
8. ———. Social schemas and the reconstruction of social object displays from memory. *J. Abn. & Soc. Psychol.*, 1962, **65**, 71-74.
9. ———. The pervasive influence of social schemata. *J. Abn. & Soc. Psychol.*, 1964, **68**, 248-254.
10. KUETHE, J. L., & STRICKER, G. Man and woman: Social schemata of males and females. *Psychol. Rep.*, 1963, **13**, 655-661.
11. LITTLE, K. B. Personal space. *J. Exper. Soc. Psychol.*, 1965, **1**, 237-247.
12. LONG, B. H., ZILLER, R. C., & HENDERSON, E. H. Developmental changes in the self-concept during adolescence. *School Rev.*, 1968, **76**, 210-230.
13. MCBRIDE, G., KING, M. G., & JAMES, J. W. Social proximity effects on galvanic skin responses in adult humans. *J. of Psychol.*, 1965, **61**, 153-157.
14. SOMMER, R. Studies in personal space. *Sociometry*, 1959, **22**, 247-260.
15. WINER, B. J. *Statistical Principles in Experimental Design* (2nd ed.). New York: McGraw-Hill, 1971.

Department of Psychology  
Brigham Young University  
Provo, Utah 84601





## A COMPARISON OF TWO KINDS OF DOGMATISM SCORES: ROKEACH CATEGORIES *VERSUS* OPEN-ENDED RESPONSES\*

*Rutgers-The State University*

MARION STEININGER

### SUMMARY

College students responded to 11 Dogmatism Scale items both with the categories used by Rokeach ("agree very much" to "disagree very much") and with open-ended comments. The latter were coded and scored, and the two sets of dogmatism scores were then correlated with each other and with attitudes hypothesized to be related to dogmatism. The Rokeach and coded dogmatism scores showed equally low but significant correlations with attitudes. The dogmatism scores were significantly correlated with each other; the students' open-ended responses fit Rokeach's theory well and did not suggest any new aspects of dogmatism. The data provide further evidence for the construct validity of Rokeach's scale.

### A. INTRODUCTION

The scoring that Rokeach devised for his Dogmatism Scale (2) involves summing the responses to the items, which range from +3 (agree very much) to -3 (disagree very much), with no zero category. In a recent study (4), such dogmatism scores were correlated with responses to nine attitude items chosen because they were conceptually related to dogmatism. These items dealt with future presidential preferences, Nixon, Agnew, Viet Nam, marijuana, characteristics of a good professor, test anxiety, church attendance, and reasons for attending college. The responses to these items were scored from 1 to 5, according to the degree to which they were theoretically assumed to reflect dogmatism; an attitude mean was then calculated for each student. For both male and female students, the correlations between dogmatism and attitude scores were low but significant.

One purpose of the present study was to determine whether scores based on open-ended responses to Dogmatism Scale items would correlate better with attitudes. A second purpose was to explore the construct validity of the Dogmatism Scale by correlating the two kinds of dogmatism scores (those based on Rokeach's scoring system and those based on the free responses), and by seeing whether any additional dimensions of dogmatism would be dis-

\* Received in the Editorial Office on September 28, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

covered in the free responses, aside from those described by Rokeach (2). This approach to studying the construct validity of the Dogmatism Scale is in line with Cronbach and Meehl's suggestion that studies of process are helpful in establishing the construct validity of a test (1).

Since people responded to the same Dogmatism Scale items in two different ways, it seemed probable that the two sets of dogmatism scores would be significantly correlated; however,  $r$  was not expected to be close to 1.00, since responding to items with Rokeach's response categories differs psychologically from free responding. Responding "acceptably" is probably easier for some people with Rokeach's categories, while others might censor their spontaneous comments more extensively. In view of this, it did not seem possible to hypothesize about the relative abilities of the two kinds of dogmatism scores to "predict" attitude scores.

Thus, three questions were explored: What is the correlation between two measures of dogmatism? What is the correlation of each with attitudes? Do spontaneous comments reveal any new aspects of dogmatism?

#### B. SUBJECTS AND PROCEDURE

Ninety-eight male and 79 female students in an introductory psychology course anonymously answered 11 items from the Dogmatism Scale (items 54, 5, 39, 30, 6, 12, 47, 53, 55, 41, 48, in that order) and the nine attitude items described earlier. The students were also asked to explain why they felt as they did about each Dogmatism Scale item.

A code was developed which followed closely the ideas expressed by the students. For example, one of the items used was the "free speech" item.<sup>1</sup> The code covered three ideas: Did the student defend free speech or favor its restriction? How did (s)he explain why free speech was good or bad? Did (s)he make any negative comments about groups of people (e.g., "uneducated masses who can be fooled")?<sup>2</sup> Coder agreement was 81% initially; disagreements were resolved by discussion. In a few instances, no agreement could be reached; the author then made the final decision.<sup>3</sup>

A coded dogmatism score was obtained for each student in the following manner. Code categories most clearly related to Rokeach's theory were scored "open-minded" or "closed-minded." Thus, the response, "Some groups

<sup>1</sup> "Even though freedom of speech for all groups is a worthwhile goal, it is unfortunately necessary to restrict the freedom of certain political groups."

<sup>2</sup> The complete code is available from the author upon request to the address shown at the end of this article.

<sup>3</sup> The author is indebted to Joanne Barraclough, Barbara Durso, and Carolyn Pasquariello for their patience and good judgment in coding.

shouldn't have freedom of speech because they don't know right from wrong," was counted as "closed-minded," while "If any group is restricted, we don't have freedom of speech any more" was counted as "open-minded." The difference between the number of "open" responses and the number of "closed" responses was used as the student's coded dogmatism score. Scores ranged from  $-7$  (high dogmatism) to  $+8$  (low dogmatism).

A combined dogmatism index was also obtained for each student by converting both his Rokeach and coded scores to 100-point scales and summing the converted scores.

### C. RESULTS

The correlations between the Rokeach and coded dogmatism scores were .61 for the men ( $p < .01$ ) and .56 for the women ( $p < .01$ ). For both measures of dogmatism, there were low but significant correlations with attitude scores. For the Rokeach scores, the  $r$ s were .34 for the men and .29 for the women; for the coded dogmatism scores, the  $r$ s were .21 and .35. The differences among these  $r$ s were not significant ( $z = 1.00$  or less). Thus, the two measures of dogmatism were significantly related, though they were clearly not identical; furthermore, neither measure was superior to the other in correlating with attitudes.

When the two measures of dogmatism were combined, the correlations between these combined scores and attitudes were .30 for the men and .36 for the women. Thus, combining the two measures of dogmatism did not improve the correlation with attitudes.

Detailed analysis showed that agreement or disagreement with an item was closely related to the content of the subsequent comments about it, but not completely. For example, about 90% of those agreeing with the "free speech" item wrote about "dangerous groups," while about 75% of those disagreeing wrote about the meaninglessness of free speech if anyone is restricted, or that everyone is equal. However, the  $+3$  to  $-3$  responses did not always correspond to the written comments in direction.

### D. DISCUSSION

In recent years, there has been renewed interest in incorporating "what the person thinks" into psychology. In the present study, the students responded in two ways to the Dogmatism Scale items: by agreeing or disagreeing to a certain degree, and by writing their thoughts about each item. The two methods were equally correlated with attitude scores. Why weren't the coded scores superior, since they reflected the respondents' spontaneous and, at times,



extensive thoughts? One asset of open-ended questions is "... the exploration of a *process* or of the individual's formulation of an issue" (3, p. 262). Especially helpful in the exploratory stages of research, open-ended responses may lead item writers to include items or response alternatives of which they would not otherwise have thought. In the present instance, however, the free responses turned out not to tap any dimension or degree of dogmatism that wasn't already included in the items. Thus, a student's comments about the free speech item may have revealed that (s)he dichotomized people by stating that there were two kinds of people, those who should have free speech and those who shouldn't; the potential gain from that comment was offset by the fact that another item tapped a very closely related idea: "There are two kinds of people in the world: those who are for the truth and those who are against the truth."

The significant correlations in two samples (male and female) between the coded and Rokeach dogmatism scores seem to provide further evidence for the construct validity of Rokeach's scale, since the free comments seem to fit Rokeach's theory well. For example, an expression of disdain for "uneducated masses" who would be fooled by those "abusing" free speech could be interpreted as encompassing one or more of the following dimensions discussed by Rokeach: self-aggrandizement, paranoid outlook, intolerance for others' views, and dichotomizing people.

Some of the discrepancies between Rokeach and open-ended responses may have occurred because the students sometimes seemed to be explaining why they didn't agree or disagree even more intensely, rather than explaining why they didn't give a response in the opposite direction. For example, a student may have disagreed "on the whole" (—2) with the free speech item, and then stated that dangerous groups had to be restricted. This occurred too often to be attributed to careless checking of response categories; rather, it seemed that the student was explaining why (s)he didn't check "disagree very much" (—3). There is no basis for stating that one type of dogmatism score measures the construct "dogmatism" better than the other, especially in view of their comparable correlations with the attitude measure; it would seem, however, that the imperfect correlation between the two measures of dogmatism should not be interpreted to mean that they are measuring different constructs.

In view of the labor involved in coding and its lack of greater validity, it seems likely that open-ended responding has no value in the use of the Dogmatism Scale. However, the correlation between the two kinds of dogmatism scores suggests the possibility that in some research, when there is concern about what effects responding to a scale might have on the dependent variable,

one could estimate whether subjects are "high" or "low" dogmatics by engaging them in a standardized conversation and counting their "open" and "closed" comments.

## REFERENCES

1. CRONBACH, L. J., & MEEHL, P. E. Construct validity in psychological tests. *Psychol. Bull.*, 1955, **52**, 281-302.
2. ROKEACH, M. *The Open and Closed Mind*. New York: Basic Books, 1960.
3. SELTZ, C., JAHODA, M., DEUTSCH, M., & COOK, S. *Research Methods in Social Relations* (rev. ed.). New York: Holt, Rinehart & Winston, 1959.
4. STEININGER, M. P., DURSO, B. E., & PASQUARIELLO, C. Dogmatism and attitudes. *Psychol. Rep.*, 1972, **30**, 151-157.

*Psychology Department*  
*Rutgers-The State University*  
*Camden, New Jersey 08102*





## A COMPARISON OF HUSBAND-WIFE RESPONSES TO PREGNANCY\*<sup>1</sup>

*Department of Psychology, University of Calgary; and Department of Psychiatry,  
Foothills Hospital, Calgary, Alberta, Canada*

---

STEWART MEIKLE AND RICHARD GERRITSE

---

### SUMMARY

Seventy married couples applying for a therapeutic abortion were asked to complete the Minnesota Multiphasic Personality Inventory (MMPI) prior to the wives' undergoing a therapeutic abortion. In addition, the wives completed a second MMPI six months after the abortion was performed. The results were interpreted as support for the view that the female bears the brunt of the psychological stress involved in coping with an unwanted pregnancy. Furthermore, the effect of such an event appears to minimize any pre-existing similarities between husbands and wives. However, the abortion procedure itself seems to some extent to redress this situation and results in an increase in similarities between spouses. The particular reaction pattern shown by the wives to stress induced by an unwanted pregnancy shows some stability over a six month period.

---

### A. INTRODUCTION

Personality similarities between husbands and wives have been extensively studied within both normal and psychiatrically disturbed groups. In the normal area earlier investigators, such as Burgess and Cotterell (1), Terman (7), and Hollingshead (4), all found evidence in support of a principle of homogamy (like marries like). More recently, however, Winch (8) has proposed an alternative viewpoint based upon a principle of heterogamy or complementary needs. To date no adequate resolution of these two positions has appeared, although presumably both factors would have to be accounted for in any comprehensive formulation. A second line of approach has involved the use of abnormal groups in which one or other of the married partners

---

\* Received in the Editorial Office on September 29, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was supported in part by National Health Grant No. 608-7-122 from the Department of National Health and Welfare.

was considered to be psychiatrically disturbed. Gregory (2), Ryle and Hamilton (6), and Kreitman (5) all report more evidence of psychological or physical symptoms among spouses of psychiatric patients than controls. The present study shared some of the characteristics of both these approaches in that it was concerned with the effects of a relatively acute life-stress situation (an unwanted pregnancy) on normal husband-wife pairs. It also considered the effects of removal of this stress. More specifically, three related but distinct questions were posed. Firstly, does an unwanted pregnancy produce the same degree of disturbance in husbands and wives? Secondly, is removal of this stress therapeutic for the female? Thirdly, when couples are exposed to this type of stress, is there any evidence that they show a greater similarity of response compared to unrelated individuals?

## B. METHOD

### 1. *Subjects*

The subjects consisted of 70 pairs of husbands and wives who applied for a therapeutic abortion at a general teaching hospital. The mean age of the husbands was 36.9 and of the wives was 33.3 years. All of the women were in the first trimester of pregnancy and already had an average of 2.8 previous children. The 70 couples were part of a larger initial group seeking abortion. One in every four of these original applications was rejected, but all 70 pairs in the present study were approved. Despite the fact that histories of medical and psychiatric difficulties were frequently presented in support of particular applications, in most instances the decision to grant or withhold was determined by current social or psychiatric factors which appeared to be mainly reactive to the unwanted pregnancy itself. It is felt, therefore, that this group of subjects were mainly normal people faced by a stressful life situation.

### 2. *Materials and Procedure*

The main dependent variables consisted of the various Minnesota Multiphasic Personality Inventory (MMPI) subscales (3). This test was administered to each couple prior to their being seen by a Therapeutic Abortion Committee, but the results were not made available to the committee to avoid possible contamination. In addition, a repeat MMPI was completed by the wives six months after the abortion procedure. The significance of the differences between the various groups was determined by *t* tests. Similarity was judged by means of Pearson product-moment correlations between subscale scores.

## C. RESULTS

Table 1 below summarizes the mean MMPI profiles for the various groups and also indicates the level of significance of the comparisons between the various subscales. An examination of rows 1 and 2 indicates that prior to the abortion procedure the wives' MMPIs are considerably more elevated than the husbands'. Row 5 shows that 11 of these 13 scale differences reach significance. It appears, therefore, that the effects of an unwanted pregnancy are more strongly felt by the wife than the husband. The husbands' profiles are in fact almost within normal limits apart from a moderate elevation on the depression scale. When the wives' six month postabortion profile is compared with their preabortion one, it can be seen that there has been a general drop in elevation. Row 4 shows that the reduction in scores reaches significance on 10 of the 13 scales. At this point, as row 6 shows, the number of wives' scales significantly elevated above their husbands' has dropped from 11 to three, and only one of these, Depression, remains above a *t* score of 60. On one scale, Hypomania, the wives are now at a significantly lower level than their husbands. In other words, six months after the removal of the stress caused by unwanted pregnancy, the wives' test profiles are virtually indistinguishable from those of their husbands obtained prior to the abortion. In general then it would appear that in the present instance the abortion procedure proved to be therapeutic.

The extent to which the husbands and wives responded to the stress of unwanted pregnancy in a similar manner was measured by running correlations between the couple's scores on the individual subscales. In addition, the women's pre- and postabortion scores were correlated. Table 2 below summarizes these results.

Examination of the first row, which correlates husbands' and wives' scores prior to abortion, indicates very little concordance. None of the correlations on the 10 clinical scales reaches significance. Only on Lie (L) and Validity (F) is any similarity revealed, and this mainly reflects test taking attitude insofar as couples tended to be somewhat similar in the degree of frankness with which they approached the test. This pattern changes, however, six months after the abortion. At this point the wives' postabortion profiles show an increase in the number of similarities compared to their husbands' tests taken prior to the abortion. This change is reflected in row 2 of Table 2 in which the number of significant correlations has now increased from two to five. Furthermore, three of these occur on the clinical scales of Psychopathic Deviate, Psychasthenia, and Social Introversion, and so are more psychiat-



TABLE 1  
MMPI *t* SCORES AND SIGNIFICANCE LEVELS BETWEEN GROUPS

Comparison groups	L	F	K	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma	Si
Standard <i>t</i> scores													
Wives preabortion	51	63	50	66	75	71	66	49	66	69	70	56	64
Husbands preabortion	49	57	52	53	60	57	58	58	54	54	54	57	52
Wives postabortion	51	57	54	55	64	60	59	48	56	56	56	49	55
Significance levels													
Wives preabortion <i>vs.</i> wives postabortion	.646	.016	.086	.001	.001	.001	.020	.568	.001	.001	.001	.004	.006
Wives preabortion <i>vs.</i> husbands preabortion	.190	.006	.154	.001	.001	.001	.001	.001	.001	.001	.001	.596	.001
Wives postabortion <i>vs.</i> husbands preabortion	.082	.818	.272	.116	.012	.116	.472	.001	.068	.030	.078	.001	.072

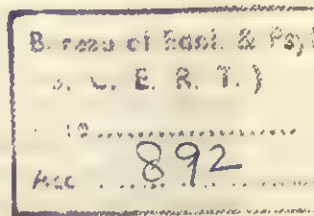
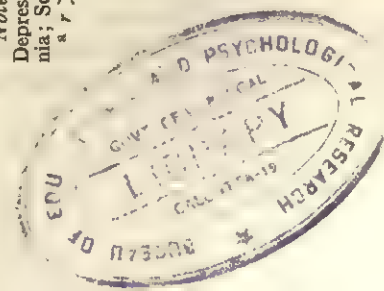
*Note:* (Definition of abbreviations) L = Lie; F = Validity; K = Test Taking Attitude; Hs = Hypochondriasis; D = Depression; Hy = Hysteria; Pd = Psychopathic Deviate; Mf = Masculinity-Femininity; Pa = Paranoia; Pt = Psychasthenia; Sc = Schizophrenia; Ma = Hypomania; Si = Social Introversion.

TABLE 2  
CORRELATIONS BETWEEN GROUPS<sup>a</sup>

Comparison groups	L	F	K	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma	Si
Wives preabortion vs. husbands preabortion	.37	.28	.22	.17	.19	.07	.13	.20	.15	-.02	.09	.01	.23
Wives postabortion vs. husbands preabortion	.09	.56	.42	.15	.23	.09	.31	.17	.12	.30	.25	.05	.29
Wives preabortion vs. wives postabortion	.58	.75	.48	.46	.46	.26	.52	.56	.44	.38	.52	.58	.67

Note: (Definition of abbreviations) L = Lie; F = Validity; K = Test Taking Attitude; Hs = Hypochondriasis; D = Depression; Hy = Hysteria; Pd = Psychopathic Deviate; Mf = Masculinity-Femininity; Pa = Paranoia; Pt = Psychasthenia; Sc = Schizophrenia; Ma = Hypomania; Si = Social Introversion.

<sup>a</sup>  $r > .35$ , significant at the .01 level;  $r > .27$ , significant at the .05 level.



rically relevant. This suggests that with the acute stress of unwanted pregnancy removed, the wives and husbands have more personality characteristics in common.

Row 3 of Table 2 contains the correlation coefficients between the wives' scores prior to abortion and their retests six months later. These are all positively significant and range from .75 for the F Scale to .26 for Hysteria. This relationship holds despite the fact that on retest the wives' scores have dropped to close to normal. In brief, then, despite general reduction in pathology, the number of symptoms a woman experienced prior to the abortion is still a moderately good predictor of the number she reports six months later.

For comparison of the wives' postabortion MMPIs with those of their husbands, it would have been more satisfactory to have had the husbands complete a second test at the same time, also. However, for a variety of reasons, this proved impractical.

#### D. DISCUSSION

It appears from the findings reported above that when couples are faced with the threat of an unwanted pregnancy, the consequent stress is more clearly reflected in the wives' test results than in the husbands'. From one point of view this result is hardly surprising, yet on the other hand, one might have expected husbands living in close proximity to relatively disturbed females to show at least some secondary emotional upheaval. The possibility was considered that either the females might have exaggerated their difficulties in order to maximize the chance of abortion, or that the males might have suppressed their true reactions; however, the MMPI validity patterns of both groups rule out the likelihood of distortion from this source seriously influencing the results. To the authors a more likely explanation for the differences between the sexes revolves around the fact that the physical changes involved in pregnancy provide immediate and continuous reminders to the woman of the true state of affairs. In addition, the occupational patterns of most husbands allow them to escape from the disturbed home situation for eight or more hours a day, thus providing a source of relief from the stress. Distractions of this sort were not so freely available to the majority of the wives in the present study.

In discussing the various sets of correlations obtained, it is perhaps worth stressing that although a number of these are significant, they are of modest magnitude. This is particularly true in the case of the two husband-wife sets. In other words, the husbands and wives are distinguished more by their MMPI differences than their similarities. However, even when this is granted,



the fact remains that there were fewer husband-wife similarities prior to abortion than after it. In other words, it appears that the effect of an unwanted pregnancy on the wives was sufficiently great to swamp whatever personality similarities they had previously shared with their husbands. The implication here is that far from an unwanted pregnancy drawing husband and wife together, it seems more likely to put distance between them, at least in terms of reducing their personality similarities. The fact that comparison of the husbands' preabortion scores with wives' postabortion scores showed an increased number of significant correlations is felt to support the view that therapeutic abortions generally decrease the differences between husbands and wives brought on by the unwanted pregnancy. If similarity between husbands and wives is considered a desirable end, then the present results favor the view that, in this instance at least, abortion was therapeutic.

Finally, the fact that the test-retest correlations between the wives' two testings were all significantly positive indicates a certain degree of stability in their presentation of themselves over the six month period. Although the characteristics that distinguished a woman prior to the abortion procedure were less obvious after the abortion as a result of a general drop in the mean profile, nevertheless the unique features of her original test pattern were still to some extent evident. Thus the wife who was reacting to the unwanted pregnancy by excessive social withdrawal (high Social Introversion score) tended to remain relatively high on this characteristic six months later. This could be interpreted as suggesting that the personality characteristics most distinctive of any particular female become exaggerated in the face of stress.

#### REFERENCES

1. BURGESS, E. W., & COTTRELL, L. *Promoting Success or Failure in Marriage*. New York: Prentice-Hall, 1939.
2. GREGORY, I. Husbands and wives admitted to mental hospital. *J. Ment. Sci.*, 1959, **105**, 457.
3. HATHAWAY, S. R., & MCKINLEY, J. C. *Minnesota Multiphasic Personality Inventory, Manual*. New York: Psychol. Corp., 1967.
4. HOLLINGSHEAD, A. B. Cultural factors in the selection of marriage mates. *Amer. Sociolog. Rev.*, 1950, **15**, 619-627.
5. KREITMAN, N. The patient's spouse. *Brit. J. Psychiat.*, 1964, **110**, 159-173.
6. RYLE, A., & HAMILTON, M. Neurosis in 50 married couples. *J. Ment. Sci.*, 1962, **108**, 265.
7. TERMAN, L. M. *Psychological Factors in Marital Happiness*. New York: McGraw-Hill, 1938.
8. WINCH, R. F. *Mate Selection: A Study of Complementary Needs*. New York: Harper, 1958.

*Department of Psychology*  
*University of Calgary*  
*Calgary, Alberta, Canada*



## CREATIVITY: PERFORMANCE, PROFILES, AND PERCEPTIONS\*

*Department of Educational Psychology, The University of Georgia*

---

VICTOR K. PHILLIPS

---

### SUMMARY

The present study investigated the relationship between creative performance, personality profiles, and self-descriptions of 100 undergraduate students. Ss were divided by a median split on the Torrance Tests of Creative Thinking (Verbal Form B) into high and low creative groups, and compared with respect to personality patterns and self-perception variables. The results indicate that the high and low creative Ss differed significantly in terms of personality factors and the way in which they perceived themselves. Also, within each group significant relationships existed between personality profiles and self-perceptions.

---

### A. INTRODUCTION

One goal of research in creativity is to provide knowledge about creative behavior so as to understand better a person's individuality. Traditional attempts to gain this perspective have involved two basic methods. For the sake of convenience we shall refer to the two approaches as Type A and Type B. Type A is characterized by use of a standardized personality inventory containing combinations of scales which have been cross-validated on successful writers, musicians, and artists. Such an inventory is administered to a sample of subjects, and consequent combinations or patterns of personality dimensions are used to define creativity levels within the group. Type B is characterized by selection of a standardized battery designed to measure creative performance. Such a measure is administered to a sample, and although critical affect variables come into play, the heaviest premium is placed upon productivity.

Type A and Type B represent, in a broad sense, measures of creative feeling and thinking states: feeling, in the sense that personality dimensions purport

---

\* Received in the Editorial Office on September 29, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.



an emotional agreement between the creative examinee and a specific set of items; thinking, to the extent that generating new products requires cognitive discipline and manipulation of concepts.

But there is a third approach which should be adopted. Type C may be considered a description of creative behavior that the individual discovers or knows of himself; that is, how accurately the creative person perceives himself in relation to his personality profile and his ability to produce. The literature does not provide answers as to whether creative self-perceptions coincide with Type A and Type B assessment data. The purpose of the present study is to define creative performance methodologically and to observe within each performance group the relation between self-perceptions and personality dimensions. In doing so, Type A, Type B, and Type C methods are brought together for a keener picture of creative individuality.

## B. METHODS

### 1. *Subjects*

The Ss were 100 University of Georgia undergraduates enrolled in three educational psychology classes during the fall quarter of 1971. Ss included 24 males and 76 females. The mean age of the Ss was 21.24 years with a class level mean of junior.

### 2. *Instruments*

Creativity level for each S was obtained through the use of the Torrance Tests of Creative Thinking, Verbal Form B (3). This battery consists of seven activities: asking questions, guessing causes, guessing consequences, product improvement, unusual uses, unusual questions, and just suppose. Interscorer reliabilities were in excess of .90. Reliability and validity information (3) rendered the battery adequate for use in the present study.

Measures of creative self-description were provided by use of the What Kind of Person Are You? checklist (4). This instrument consists of 50 items, each of which contains two behavioral descriptions. Each set of alternatives is designed to differentiate between high and low creative individuals. Correct items are those corresponding to highly creative characteristics. Thus, the correct items are summed to indicate the creative propensities of the examinee through self-description. The reliability and validity evidence for the checklist was adequate for use in the present research.

Personality information for each S was gained by administration of the Omnibus Personality Inventory, Form F (2). This Inventory contains 385

true-false items which yield separate scores for 14 personality dimensions. The dimensions are the following: thinking introversion (TI), theoretical orientation (TO), estheticism (Es), complexity (Co), autonomy (Au), religious orientation (RO), social extroversion (SE), impulse expression (IE), personal integration (PI), anxiety level (AL), altruism (Am), practical outlook (PO), masculinity-femininity (MF), and response bias (RB). Reliability and validity material is extensively presented in the norms manual (2).

### 3. *Data Collection*

Data were collected over a period of two weeks (November, 1971). The What Kind of Person Are You? (WKPAY) was administered in groups to each of the three classes. Each group was given standard instructions and allowed unlimited time. All WKPAY forms were scored according to the standard scoring key. A single score was derived for each subject.

The Torrance Tests of Creative Thinking, Verbal Form B (TTCT), was given in groups one week after the WKPAY was administered. Each group was allowed only the specified time limits for each activity. The booklets were scored according to the published scoring guide. Scoring provided measures of fluency, flexibility, and originality for each student. The component scores were then summed in order to obtain an index reflecting the total creative performance for each S.

The Omnibus Personality Inventory, Form F (OPI), was administered to each group approximately one week after the TTCT. Standard instructions were read aloud, and each group was allowed unlimited time in which to complete the 385 items.

### C. RESULTS

All Ss were ranked from lowest to highest on the TTCT measures, and a median split defined high and low creative groups. Table 1 contains the mean and standard deviations of the raw scores for the two groups and the tests of significance.

With the exception of the PO and MF scales, the high creative performance group had larger means than the low group. Of the comparisons between high and low creative groups, seven scales differed significantly to at least the .06 level, while the remaining seven were found not to be significantly different, as shown in Table 1. High scores on the TI, Es, Au, RO, and Am scales indicate, respectively, a tendency to examine their own motives and reactions; enjoyment of poetry, paintings, architecture, and reading about artistic and literary achievements; higher self-actualization and rebelliousness; more

TABLE 1  
MEANS, SDs, AND TESTS OF SIGNIFICANCE FOR *What Kind of a Person Are You?*  
AND *Omnibus Personality Inventory* MEASURES ACROSS HIGH AND LOW  
CREATIVITY DEFINED BY MEDIAN SPLIT ON *Torrance Tests of*  
*Creative Thinking* SCORES

Measure	TTCT high ( <i>N</i> = 50)		TTCT low ( <i>N</i> = 50)		<i>F</i>	<i>p</i>
	Mean	<i>SD</i>	Mean	<i>SD</i>		
WKPAY	25.02	6.61	22.70	6.48	3.14	<.07
Thinking introversion (TI)	24.04	6.92	20.22	8.13	6.39	<.02
Theoretical orientation (TO)	15.48	5.02	13.86	5.85	2.20	NS
Estheticism (Es)	13.84	5.16	11.56	5.28	4.77	<.03
Complexity (Co)	15.52	5.20	13.80	5.63	2.34	NS
Autonomy (Au)	28.02	7.59	24.10	7.22	7.00	<.01
Religious orientation (RO)	12.48	5.53	10.40	5.46	3.58	<.06
Social extroversion (SE)	22.70	7.34	20.60	7.19	2.08	NS
Impulse expression (IE)	30.34	9.42	28.12	8.13	1.58	NS
Personal integration (PI)	31.44	8.78	30.64	12.05	0.14	NS
Anxiety level (AL)	11.64	4.19	11.42	5.09	0.05	NS
Altruism (Am)	22.08	6.00	18.96	6.93	5.78	<.02
Practical outlook (PO)	13.64	5.03	16.32	5.81	6.06	<.02
Masculinity-femininity (MF)	22.94	6.02	25.60	6.27	4.67	<.03
Response bias (RB)	10.76	3.51	10.48	4.39	0.12	NS

liberal views about the existence of God; and higher affiliation needs. Low scores on the same dimensions reflect, respectively, a dislike for philosophical or serious books and a greater need for environmental certitude; a lack of interest in the fine arts or artistic things; a high degree of need for structure and controlled guidance; a basic and rather unshakeable belief in God; and a lack of responsibility.

Table 2 shows the relationships between the personality dimensions and the self-descriptions within the two creativity levels and within the total sample. The coefficients of correlation within the total sample are significant at the .01 level for all of the scales identified by Heist (1) as characteristic of highly creative college students. These are TI, TO, Es, Co, Au, RO, and IE. In addition PO is negatively correlated at the .01 level, and SE, PI, and MF are negatively correlated at the .05 level. Similar patterns are noted within the two creativity levels. Perhaps because of the restriction of range, the level of significance of some of the correlation coefficients within creativity levels drops.

In terms of OPI profiles, subjects scoring high on the WKPAY instrument may be characterized as liking reflective thought and academic activities; expressing interest in a wide range of ideas in a variety of areas; preferring to deal with theoretical concerns and problems and to use the scientific method in thinking; expressing interests in artistic matters and a high level of sensitivity and response to esthetic stimulation; having a high tolerance for



TABLE 2  
PRODUCT-MOMENT CORRELATION COEFFICIENTS AND RESPECTIVE SIGNIFICANCE LEVELS  
BETWEEN THE *Omnibus Personality Inventory* and TOTAL SAMPLE AND  
*What Kind of a Person Are You?* RAW SCORES FOR  
HIGH AND LOW CREATIVITY GROUPS

OPI	TTCT high (N = 50)		TTCT low (N = 50)		Total (N = 100)	
	r	p	r	p	r	p
Thinking introversion (TI)	.38	<.01	.39	<.01	.45	<.01
Theoretical orientation (TO)	.56	<.01	.44	<.01	.52	<.01
Estheticism (Es)	.43	<.01	.57	<.01	.48	<.01
Complexity (Co)	.56	<.01	.53	<.01	.54	<.01
Autonomy (Au)	.16	NS	.57	<.01	.38	<.01
Religious orientation (RO)	.28	<.05	.33	<.01	.32	<.01
Social extroversion (SE)	-.28	<.05	-.16	NS	-.19	<.05
Impulse expression (IE)	.31	<.05	.40	<.01	.36	<.01
Personal integration (PI)	-.10	NS	-.19	NS	-.19	<.05
Anxiety level (AL)	-.02	NS	-.21	NS	-.11	NS
Altruism (Am)	-.06	NS	-.14	NS	-.06	NS
Practical outlook (PO)	-.37	<.01	-.54	<.01	-.49	<.01
Masculinity-femininity (MF)	.01	NS	-.33	<.01	-.20	<.05
Response bias (RB)	.00	NS	.06	NS	.13	NS

Note: TTCT = Torrance Tests of Creative Thinking, Verbal Form B.

ambiguities and uncertainties and liking for novel and complex situations and ideas; tending to be independent of authority and opposing infringements on the rights of individuals; manifesting a moderate view of religious beliefs and practices; having an active imagination, expressing impulses, valuing sensual reactions and feelings, finding a greater appeal in ideas than in facts.

#### D. DISCUSSION

The findings reported here give evidence that the highly creative person perceives himself fairly accurately in terms of Type A and Type B assessment data. That is, if he is creative, he will not only perform creatively, but will manifest a creative personality profile and describe himself in terms that have characterized recognized creative behavior. Of equal interest is the accuracy with which the low creative individual sees himself. Chances are, if he is unable to generate original productions, his personality profile will convey a rather rigid, inflexible, and conforming pattern given to tradition. But, at the same time, he will describe his less creative behavior with a significant degree of accuracy.

It should also be noted that the results reported herein add to the construct validity of both the verbal form of the Torrance Tests of Creative Thinking and the *What Kind of Person Are You?* instrument when used with college students.

## REFERENCES

1. HEIST, P., Ed. *The Creative College Student: An Unmet Challenge*. San Francisco: Jossey-Bass, 1970.
2. HEIST, P., & YONGE, G. *Omnibus Personality Inventory, Form F.: Manual*. New York: Psychol. Corp., 1968.
3. TORRANCE, E. P. *Torrance Tests of Creative Thinking: Norms-Technical Manual* (res. ed.). Princeton, N. J.: Personnel Press, 1966.
4. ———. *What Kind of Person Are You? Technical Norms Manual* (res. ed., mimeo.). Athens, Ga.: Georgia Stud. Creative Behav., 1970.

*Department of Educational Psychology*

*The University of Georgia*

*Athens, Georgia 30601*

## AWARENESS IN VERBAL OPERANT CONDITIONING: EXAMINATION OF PERFORMANCE QUALITY CHANGES\*

*University of Missouri-Columbia*

---

EVERETT E. ADAM, JR., AND ROBERT J. PAUL

---

### SUMMARY

The hypothesis addressed in this study was that for every experimental condition, learning with awareness would occur. An experiment was conducted with use of a collating task that incorporated the awareness question into the design investigating the change in performance quality and quantity. Statistical evaluation of Dulany's awareness questionnaire suggested general unawareness of attempted verbal operant conditioning. There is some evidence suggesting that an actual *change* in performance quality and quantity is accompanied by awareness with use of operant conditioning procedures.

---

### A. INTRODUCTION

The basic issue of whether subjects are in fact aware or unaware when they learn through verbal operant conditioning procedures appears unresolved. Literature reviews present experimental results and the issue in detail, as exemplified by Adams (3), Postman and Sassenrath (10), DeNike (5), Williams (16), Spielberger and DeNike (13), and Thaver and Oaks (14). The position for unawareness is typified by Postman and Sassenrath (10), Marlowe (9), and Thaver and Oaks (14). A rather rigid definition of awareness is accepted by most researchers supporting learning without awareness. The other stand, that supporting awareness in verbal operant conditioning, is exemplified by Adams (3), Dulany (6), Spielberger (12), and DeNike (5). Of particular interest here is the procedural improvement by DeNike when he had subjects write their responses indicating awareness or unawareness.

Hersen (8) suggests that awareness is a function of many intervening variables, some of which are the point at which the interview is administered, the complexity of the response class conditioned, and the interview specificity. Verplanck (15, p. 130) states quite clearly that the awareness dispute is a

---

\* Received in the Editorial Office on October 4, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

regrettable one, as it seems to have led to the performance of experiments on inappropriate forms of behavior and to a proliferation of speculative theory. He states:

By inappropriate forms of behavior, I mean this: The experiments that have been—by now—repeated over and over with only minor modifications are those that have confounded at least two questions, the identification of response classes and the stability (habitability) of reinforcers. "Saying plural nouns," "constructing sentences in the first person," "Mm-hmm," and "Good" may serve to demonstrate the occurrence of operant conditioning, but they are not necessarily the best choices for experiments on other problems (15, p. 130).

In consideration of the above, a completely different problem was chosen for this study. The basic study was designed to investigate the applicability of operant conditioning to the organizational problem of motivation for improved performance quality. This operant conditioning, quality motivation viewpoint was formulated by Adam and Scott (2) and then experimentally tested by Adam (1). It is in this latter study that verbal operant conditioning was attempted. Therefore, the purpose of this paper is to report results on the "learning without awareness" question when the problem and conditioning procedures were chosen to represent a real organizational problem—the problem of influencing performance quality on routine repetitive tasks often found in industrial settings. The hypothesis to be tested is that for every experimental condition, learning with awareness took place. Further, if there was no learning, it is hypothesized that unawareness of the response-reinforcement contingency would be reported.

## B. METHOD

### 1. *Subjects and Apparatus*

The 80 subjects who participated in the study were males from junior- and senior-level undergraduate courses. Each subject worked in an individual room at a task that involved collating data processing unit record cards that were collected periodically and computer evaluated.

### 2. *Procedure*

The task involved collating six punched and interpreted data processing unit record cards, one from each of six boxes. Approximately four percent of the total cards were error cards, a typical percent defective in industrial situations, which were randomly introduced into the boxes after being punched and



coded also. The subject was instructed to remove the error cards, the cards mispunched, as he collated from the six boxes. Six collated cards completed one unit of quantity. Two types of error were possible: a failure to screen out a bad card and a sequencing error. The two errors combined into a total error count subsequently referred to as errors.

Initially, the subject was given general instructions on task performance and a statement that this was a study to obtain his feelings concerning a card sorting task about which he would be asked some questions at the end of the session. Then special instructions were given explaining the verbal reporting of performance. These instructions depended upon whether quantity or quality was to be emphasized during the first time period. The subject was told he would be given periodic reports concerning his performance and was reinforced according to Table 1 for the three-hour experimental session. Standards for both quantity and quality were established by traditional industrial engineering techniques, and reinforcement was administered according to an intermittent schedule with a fixed interval of 15 minutes between reinforcements. The conditioning literature supports a variable schedule as the better type of reinforcement schedule for maintaining performance levels over long periods of time, but a fixed interval reinforcement schedule was chosen because that is the way many reinforcers in industry are administered. Examples are a weekly pay check, an annual pay increase, and a six-month performance appraisal.

If quantity was emphasized initially, the subject was instructed, "Do the best that you can. We have found that you should work as fast as you can on this job." If quality was initially stressed, he was told, "Do the best that you can. We have found that you should try to have as few errors as you can on this job." Midway through the session the contingency was shifted for those groups that were to be changed from emphasizing quantity to quality (or from emphasizing quality to quantity). The subject was *not* openly told of the contingency shift, but was expected to infer the shift through the change in the response-reinforcement contingency. Forty of the subjects received this shift (20 from quality to quantity, and 20 from quantity to quality), and for the purposes of experimental control 40 underwent no contingency shift (20 quality throughout, and 20 quantity throughout). In order to maintain the response-reinforcement contingency, the reinforcer depended solely upon the experimental treatment and the *output response*. At the end of the session the subjects were administered the questionnaire developed by Dulany (7) with the addition of question four to clarify the behavioral hypothesis for groups where the reinforcement procedure changed (see Table 2).

TABLE 1  
MATRIX OF VERBAL REINFORCER

Response reinforced	Low error and high quantity <sup>a</sup>	Subject performance		
		Low error and high quantity <sup>b</sup>	High error and high quantity <sup>c</sup>	High error and low quantity <sup>d</sup>
Quantity	(A) Very good	Acceptable	Good	Not good
	(B) Very high; you are well above the average number of units	Not high; you are below the average number of units	Very high; you are well above the average number of units	Not high; you are below the average number of units
	(C) Very good	Good	Acceptable	Not good
	(D) Very low; you are well below the average number of errors	Very low; you are well below the average number of errors	Not low; you are above the average number of errors	Not low; you are above the average number of errors

<sup>a</sup> Zero, one, or two errors; 63 or more units.

<sup>b</sup> Zero, one, or two errors; zero to 62 units.

<sup>c</sup> Three or more errors; 63 or more units.

<sup>d</sup> Three or more errors; zero to 62 units.

TABLE 2  
QUESTIONS ABOUT THE TASK

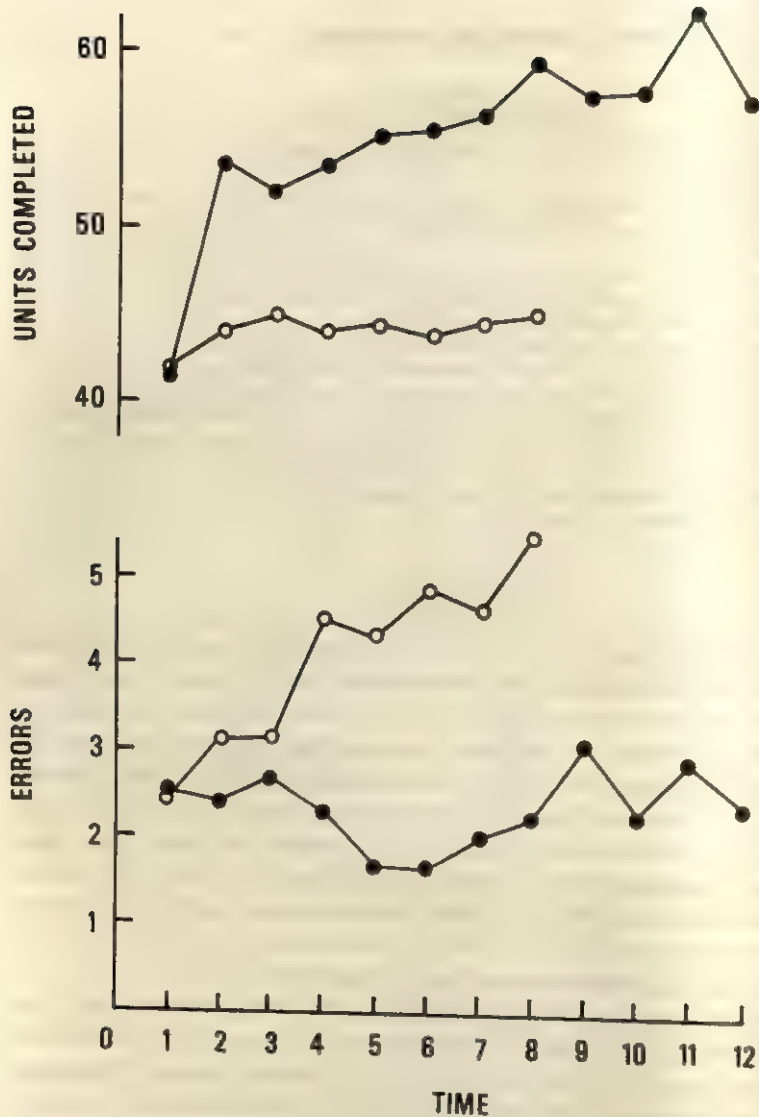
1. Did you notice whether I said anything during the study? Yes \_\_\_\_ No \_\_\_\_ If yes, what did I say?
2. Did you come to think what I said was random or did it seem to follow anything you did? Random \_\_\_\_ Followed something \_\_\_\_ If it followed something you did, what was that?
3. Did you come to think there was or wasn't any purpose to the words "acceptable," "good," and "very good"? Was a purpose \_\_\_\_ Was no purpose \_\_\_\_ If you thought there was a purpose, what was that?
4. How much would you say you wanted me to say "good" or "very good"? Very much \_\_\_\_ Some \_\_\_\_ Didn't care one way or the other \_\_\_\_
5. Did you come to think there was anything you were supposed to do in each time period in order to be correct? Yes \_\_\_\_ No \_\_\_\_ If yes, what were you to do?
6. Did you come to think that what you were supposed to do in each time period ever changed? Yes \_\_\_\_ No \_\_\_\_ If yes, what was the change?
7. In performing this task how did you go about deciding whether to work for low errors, high quantity, or both?

### C. RESULTS

Ruch (11) used this same task and sampled from essentially the same population, but he employed no reinforcement in his shorter experimental sessions. Figure 1 is representative of the various conditions, and compares verbal reinforcement to the absence of reinforcement for this task when the response-reinforcement contingency focuses on quality reinforcement throughout. The graph suggests that verbal conditioning did take place. Analysis of the difference in responses made early in the current study with those made later in the current study showed significant differences in quantity ( $p \leq .05$ ), but no differences in quality. The absence in the current study of a difference in quality should be contrasted *in direction* with Ruch's nonsignificant ( $.05 \leq p \leq .10$ ) *increase* in errors between time periods 1 and 8.

Independent ratings of subjects' responses to the seven questions (see Table 2) were made by three selected, carefully trained raters. The ratings were recorded corresponding to the three Dulany (7) response category definitions: report of reinforcement hypothesis, report of behavioral hypothesis, and report of behavioral intentions. As assigned weighting of 1 through 4 or 1 through 5 for scoring was made in accordance with the Dulany response categories. The weight of 1 corresponds to his A or 1 in his various categories. The data were then analyzed for differences between means by use of an analysis of variance randomized block two-variable classification without replication (4, 17).

Table 3 reports the mean score for each Dulany response (as indicated in the preceding paragraph), the standard deviation, and the *F* ratio for each



● QUALITY REINFORCED EACH TRIAL (CURRENT STUDY)  
 ○ NO REINFORCEMENT PROCEDURE APPLIED (RUCH'S STUDY)

FIGURE 1  
 A COMPARISON OF PERFORMANCE QUANTITY AND QUALITY UNDER VERBAL  
 REINFORCEMENT AND NO REINFORCEMENT



TABLE 3  
EXPERIMENTAL CONDITION DIFFERENCES

Question number	Reinforcement conditions	Mean	SD	F ratio
1	1	3.10	.64	2.12
	2	2.75	.55	
	3	3.10	.45	
	4	3.05	.39	
2	1	2.85	.67	2.88
	2	2.65	.59	
	3	2.85	.76	
	4	3.25	.64	
3	1	2.85	.93	1.01
	2	2.95	.89	
	3	3.25	.85	
	4	2.85	.67	
4	1	2.60	.99	3.97*
	2	3.20	.77	
	3	3.10	.71	
	4	3.45	.68	
5	1	4.70	.80	.50
	2	4.55	.51	
	3	4.50	.69	
	4	4.45	.69	
6	1	4.80	.77	4.34*
	2	4.25	.79	
	3	4.35	.76	
	4	4.95	.60	
7	1	4.65	.75	6.84**
	2	4.15	.59	
	3	4.60	.75	
	4	5.20	.83	

\* Significant at .05 level.

\*\* Significant at .01 level.

reinforcement condition: condition 1 being quality reinforced throughout; condition 2 being quality reinforced initially and then quantity reinforced subsequently; condition 3 being quantity reinforced throughout; and condition 4 being quantity reinforced initially followed by a shift to quality reinforcement. Results indicate that questions 4 and 6 are significant at the .05 level, and question 7 at the .01 level. Also of importance is the actual interpretation of mean scores with regard to the awareness issue. This will be discussed in the next section. A check for a difference between raters resulted in no differences except on reinforcement condition 2 (.05 level).

#### D. DISCUSSION AND CONCLUSIONS

The hypothesis of interest was that for every experimental condition, learning with awareness took place. It was suggested by examination of quality and

quantity output for this study, compared to Ruch's study where no reinforcement existed, that learning appeared to take place in this study (Figure 1). Statistical comparison of periods 1 and 8 within each study suggested a stronger possibility of learning in the current study than in Ruch's study.

Responses from the Dulany awareness questionnaire were made within three general categories. The responses concerning the subjects' reinforcement hypothesis slightly indicated awareness. However, on both the behavioral hypothesis and the behavioral intentions categories the subject reported unawareness. In total, the results of this study suggest unawareness in attempted verbal operant conditioning (Table 3).

Examination of each question by experimental conditions indicated that questions 4, 6, and 7 were significant between conditions, and the other four questions were not significantly different between experimental conditions. No apparent reason could be found for these particular questions showing statistical significance. Closer study of each question that was significant resulted in no further patterns for question 4. It was noted that for questions 6 and 7, condition 2 reported "most aware" and condition 4 "least aware" in each question. Condition 2 (quality primarily reinforced, shifting to quantity primarily reinforced) was reported by Adam (1) to elicit a behavioral change, while he reported that condition 4 (quantity primarily reinforced throughout) resulted in no reported behavioral change. Pairing Adam's results with the results of the awareness questionnaire, we notice that subjects reported awareness on questions 6 and 7 for the experimental condition in which they actually underwent a change in performance quality and quantity, but reported unawareness in the condition where no actual behavioral change was attained. This finding is different from the general finding of this study supporting unawareness, but might have important implications of those in organizational settings attempting to change performance quality and quantity. This suggests that when *changing* performance, one might give extra attention to attempting to make the response-reinforcement contingency well understood by the recipient: understood to the point that he is aware of a change and can see the results of the shift in the contingency relationship.

An examination of rater consistency among conditions indicated no difference between raters, except on condition 2. A further examination of rater differences by question suggests that when there were differences, they appeared to be associated most frequently with rater A. Finding rater consistency strengthens the general finding of unawareness.

In conclusion, it was found that verbal operant conditioning procedures did

not, for both quantity and quality, statistically result in learning on this representative organizational problem. This complicates considerably the ability to generalize concerning the results. Dulany's questionnaire was administered, and results indicated general unawareness of attempted verbal operant conditioning. Closer examination revealed that on some questions, when the attempt was successful in changing from quality to quantity emphasis, subjects reported significantly "higher awareness" than they did when change from quantity to quality was emphasized. In the later case, no overt change was found to be significant. Subjects reported for this condition that they were "more unaware" than for any other condition. Therefore, there is limited evidence suggesting that an actual *change* in performance quality and quantity is accompanied by awareness by use of operant conditioning procedures.

This study has provided definite changes in methodology, among them being extension of the awareness in verbal operant conditioning question to a more realistic task, and incorporation of the awareness question into an experimental design investigating the change in performance quality and quantity. A further extension was the statistical analysis of Dulany's questionnaire, thus strengthening the inferences that could be made concerning awareness. Although this study was not designed to, and did not provide a final answer for the awareness/unawareness question, perhaps some of the changes in the methodology employed herein will stimulate further meaningful investigations.

#### REFERENCES

1. ADAM, E. E. JR. An analysis of the change in performance quality employing operant conditioning procedure. *J. Appl. Psychol.*, in press.
2. ADAM, E. E., JR., & W. E. SCOTT, JR. The application of behavioral conditioning procedures to the problems of quality control. *Acad. Manag. J.*, 1971, **14**, 175-193.
3. ADAMS, J. S. Laboratory studies in behavior without awareness. *Psychol. Bull.*, 1957, **54**, 383-405.
4. CHOU, Y-L. Statistical Analysis with Business and Statistical Applications. New York: Holt, Rinehart & Winston, 1969. P. 44.
5. DENIKE, L. D. The temporal relationship between awareness and performance in verbal conditioning. *J. Exper. Psychol.*, 1964, **68**, 521-529.
6. DULANY, D. E. Hypotheses and habits in verbal operant conditioning. *J. Abn. & Soc. Psychol.*, 1961, **63**, 251-263.
7. ———. The place of hypotheses and intentions: An analysis of verbal control in verbal conditioning. In C. W. Eriksen (Ed.), *Behaviorism and Awareness*. Durham, N. C.: Duke Univ. Press, 1962. Pp. 102-109.
8. HERSEN, M. Awareness in verbal operant conditioning: Some comments. *J. Gen. Psychol.*, 1968, **78**, 287-296.
9. MARLOWE, D. Need for social approval and the operant conditioning of meaningful verbal behavior. *J. Consult. Psychol.*, 1962, **26**, 79-83.

10. POSTMAN, L., & SASSENATH, J. The automatic action of verbal rewards and punishments. *J. Gen. Psychol.*, 1961, **65**, 109-136.
11. RUCH, W. A. The effects of changes in stimulus complexity on routine, repetitive task performance. Unpublished Doctoral dissertation, Indiana University, Bloomington, 1968.
12. SPIELBERGER, C. D. The role of awareness in verbal conditioning. In C. W. Eriksen (Ed.), *Behaviorism and Awareness*. Durham, N. C.: Duke Univ. Press, 1962. Pp. 73-101.
13. SPIELBERGER, C. D., & DENIKE, L. D. Descriptive behaviorism versus cognitive theory in verbal operant conditioning. *Psychol. Rev.*, 1966, **73**(4), 306-326.
14. THAYER, F., & OAKS, W. F. Generalization and awareness in verbal operant conditioning. *J. Personal. & Soc. Psychol.*, 1967, **6**(4), 391-399.
15. VERPLANCK, W. S. Unaware of where's awareness: Some verbal operant—Notates, monents, and notants. In C. W. Eriksen (Ed.), *Behaviorism and Awareness*. Durham, N. C.: Duke Univ. Press, 1962. Pp. 130-158.
16. WILLIAM, J. H., Conditioning and verbalization: A review. *Psychol. Bull.*, 1964, **62**, 383-393.
17. WINER, R. J. *Statistical Principles in Experimental Design*. McGraw-Hill, 1962. Pp. 82-89, 174-178.

*Department of Management*  
*University of Missouri-Columbia*  
*Columbia, Missouri 65201*



## CUE UTILIZATION PATTERNS IN STUDENT-FACULTY EVALUATION\*

*University of Illinois at Urbana-Champaign*

---

STEVEN E. PERMUT

---

### SUMMARY

Judgment by students of faculty effectiveness have relied, for the most part, on subjectively expressed responses to a series of "effectiveness traits" (cues). Simple averaging schemes have summarized the aggregate judgments, often resulting in (a) treating all cues as if they were of equal importance, and (b) inattention to the effects or measurement of individual differences of judges.

Multiple regression was used in this study to derive (empirically) individual cue utilization patterns (i.e., weights) for 10 effectiveness traits. Ss were compared to their subjectively expressed judgment scheme. Results indicate that Ss were moderately successful in expressing their actual cue utilization patterns; however, different judgment schemes were clearly observed.

---

### A. INTRODUCTION

Accountability in higher education is coming under increasing scrutiny by students, faculty, administrators, and government and private agencies (3). Research attention has been directed along almost every conceivable avenue (1, 2, 4, 6, 12). A number of evaluation strategies and techniques have been offered in the hope of gaining some degree of empirical structure for assessing educational accountability in general and faculty teaching performance in particular.

Several studies have attempted to create instruments based on a factor analytic framework. For example, Meredith (11) sought to establish the dimensions of faculty-course evaluation by factor analyzing 67 variables. He identified two factors, "instructor impact" and "instructional impact," accounting for 64 percent of the rotated variance. Holmes (9) identified four factorial dimensions of faculty assessment (quality of instructors' presentations, student-instructor interactions, degree to which students were motivated and stimulated by instructor, and the clarity of examinations). Harvey and Barker (7), on the other hand, examined students' gross *subjective* judgments,

---

\* Received in the Editorial Office on October 10, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

suggesting a representative list of 10 items (objectives clarified by the instructor, organization of course, knowledge of subject, range of interests and culture, preparation for class, skill as a lecturer, skill as discussion leader, variety in classroom technique, assignments, and ability to arouse interest).

Studies such as those cited attempt to generate empirically a set of attributes (cues)—e.g., behavioral characteristics, descriptive adjectives, attitude items—that are presumed to describe or summarize some component(s) of faculty classroom performance. Once these attributes are compiled and presented for student response, a number of researchers fail to make provision for the effect of individual differences of their subjects.

An important consequence of this oversight results in treating each descriptive attribute (e.g., "organization of course," "knowledge of subject," etc.) as if it were of equal weight or importance in assessing the overall performance of an instructor. Thus, studies may compare, for example, mean values for several instructors, followed by *t* tests of all pairs of means or some similar procedure. Unfortunately, such comparisons do not suggest that subjects may differ drastically in the way attributes are cognitively weighted and combined in the evaluation process (5).

Studies of human judgment and inferential processing in a number of different contexts, however, have confirmed the existence—and importance—of differential cue utilization. Hoffman (8), Hursch, Hammond, and Hursch (10), and others have explored the facets of inferential processing using multiple-cue probability models. The general approach has been to construct a representative series of "profiles," each containing numerically defined attributes (e.g., on a 1-to-7 scale) of the target object for which judgment is to be made. Judges examine each profile and respond by simply marking a standard response scale. The values of each attribute for all profiles so judged are treated as independent variables, and regressed across the corresponding dependent measure (i.e., the judge's own evaluation of the *i*<sup>th</sup> profile). The square of the multiple correlation—the coefficient of determination—is a measure of the precision by which a (linear) combination of the cues accounts for the variance in the judgment response. Individual regression coefficients provide a subject's cue utilization pattern.

The present study explores the application of a multiple-cue probability model (11) to student evaluations of faculty. Specifically, it was felt that (a) different cue utilization patterns do, in fact, exist for students in a given course when they are asked to evaluate teaching performance; and (b) students would be relatively inaccurate in subjectively evaluating their own cue utilization pattern.

## B. METHOD

## 1. Subjects

Six male and eight female students in a summer session survey-type course at the University of Illinois participated. Mean age was  $21\frac{1}{2}$  years; 11 different major fields of study were represented. It was felt that a naturally occurring student classroom sample, even a relatively small one, was preferable to an artificially created sample.

## 2. Test Instrument

In order to generate a series of "instructor profiles" for subjects to evaluate, 10 traits as shown in Table 1 were selected from the literature previously cited (e.g., 9, 11). Because the focus of the study was on examining student cue utilization patterns regardless of the origin of the cues being judged, no empirical trait selection procedure was used.

Subjects were told that each profile represented a hypothetical (but plausible) instructor in an undergraduate survey-type course. Each profile showed the instructor's average rating across all 10 traits on a 1 to 5 scale (where 5 = excellent, 1 = poor). Subjects were asked to "examine each of the 40 instructor profiles, and, on the basis of your own personal evaluation, indicate the 'overall effectiveness of the instructor' on the seven-point bipolar scale (effective-ineffective)." Subjects were further requested to maintain the same frame of reference throughout the task, and not to return to a profile once it had been marked.

After evaluating all profiles, subjects were asked to "think back over *all*

TABLE 1  
INSTRUCTOR EFFECTIVENESS TRAITS AND THEIR COMPOSITE EMPIRICAL  
AND SUBJECTIVE ORDER OF IMPORTANCE

Effectiveness traits	Empirically derived order	Subjectively expressed order
1. Effectiveness of speech (volume, clarity, etc.)	8	10
2. Treatment of questions asked in class	5	6
3. Ability to stimulate classroom discussion	10	7
4. Ability to present material on understandable plane	4	1
5. Daily organization of lecture material	9	9
6. Ability to make class material relevant	3	4
7. Personal interest in students	6	8
8. Interest and enthusiasm in subject	7	3
9. Apparent knowledge of subject	1	2
10. Grading (fairness, thoroughness, etc.)	2	5

TABLE 2  
REGRESSION MODELS (BETA WEIGHTS) *Vet345* SUBJECTIVE MODELS (RANK ORDERINGS)

S#	Type <sup>a</sup>	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>	X <sub>10</sub>	R <sup>2</sup>	r <sub>g</sub>
I	Beta	.16	.35	.14	.34	.13	.26	.13	.07	.30	.45	.79****	
	B-rank	.6	2	7	3	9	5	8	10	4	1		
II	S-rank	10	6	9	4	7	2	8	5	3	1		.54
	Beta	.11	.05	.03	.06	.26	.74	.24	.18	.26	.03	.90****	
III	B-rank	.6	8	9	7	3	1	4	5	2	10		
	S-rank	10	9	8	2	3	1	4	6	7	5		.43
IV	Beta	.01	.16	-.05	.12	.06	.12	.39	.11	.59	.03	.76***	
	B-rank	10	3	8	4	7	5	2	6	1	9		
V	S-rank	7	4	5	6	8	9	3	2	1	10		.64**
	Beta	.18	.15	-.03	.39	.04	.21	.06	.19	.79	.21	.87****	
VI	B-rank	.6	7	10	2	9	3	8	5	1	4		
	S-rank	9	6	10	2	3	8	7	5	1	4		.56*
VII	Beta	-.01	.20	.12	.10	.11	.58	.16	.34	.30	.13	.80****	
	B-rank	10	4	7	9	8	1	5	2	3	6		
VIII	S-rank	9	2	3	4	10	1	7	5	6	8		.54
	Beta	.05	.21	.15	.56	.31	.05	.23	.10	.32	.32	.80****	
IX	B-rank	10	6	7	1	4	9	5	8	3	2		
	S-rank	7	6	4	1	8	10	9	5	2	3		.62*
X	Beta	.29	.31	.30	.35	.18	.26	.12	.19	.18	.59	.89****	
	B-rank	5	3	4	2	9	6	10	7	8	1		
XI	S-rank	5	4	3	1	8	10	9	7	6	2		.84***



TABLE 2 (continued)

S#	Type <sup>a</sup>	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>	X <sub>10</sub>	R <sup>2</sup>	r <sub>s</sub>
VIII	Beta	-.04	.08	-.06	.88	-.11	.05	.07	.01	.57	.13	.94****	
	B-rank	9	5	7	1	4	8	6	10	2	3		
	S-rank	10	4	8	1	7	9	3	5	2	6		.66**
IX	Beta	.33	.12	-.09	-.03	.19	-.02	.19	.48	.68	.12	.88****	
	B-rank	3	6	8	9	5	10	4	2	1	7		
	S-rank	5	3	4	8	6	7	9	1	2	10		.54
X	Beta	.23	.11	-.12	.25	.10	.36	.11	.06	.50	.48	.84****	
	B-rank	5	7	6	4	9	3	8	10	1	2		
	S-rank	6	4	10	3	7	1	9	8	2	5		
XI	Beta	.21	.28	.19	.02	.06	.09	.40	.14	.24	.50	.86****	
	B-rank	5	3	6	10	9	8	2	7	4	1		
	S-rank	10	8	6	5	9	4	2	7	1	3		.37
XII	Beta	.14	.24	.02	.26	.09	.52	.17	.32	.37	.36	.83****	
	B-rank	8	6	10	5	9	1	7	4	2	3		
	S-rank	9	7	8	2	6	3	10	1	4	5		.67**
XIII	Beta	.19	.25	.11	.30	-.02	.40	.23	.02	.13	.60	.90****	
	B-rank	6	4	8	3	9	2	5	10	7	1		
	S-rank	9	7	6	1	10	4	2	5	8	3		.58*
XIV	Beta	.29	.23	.34	.23	.41	.24	.26	.27	.37	.23	.89****	
	B-rank	4	10	3	9	1	7	6	5	2	8		
	S-rank	8	5	3	2	10	4	6	7	1	9		-.13

<sup>a</sup> Type refers to: Beta = beta weights; B-rank = beat weight rank order (absolute magnitude using 3 decimal points); S-rank = subjectively expressed rank order.

\*  $p \leq .05$ .

\*\*  $p \leq .025$ .

\*\*\*  $p \leq .005$ .

\*\*\*\*  $p \leq .001$ .

profiles and try to indicate the *relative importance* of each trait in arriving at your overall effectiveness evaluations." A simple ranking procedure was used, where 1 = most important, 2 = next most important, through 10 = least important. Average completion time for the entire session was 24 minutes.

### C. RESULTS

Multiple linear regression provided beta weights (standardized regression coefficients) for each of the 10 traits across 40 profiles. Table 2 shows each subject's regression equation and the corresponding subjectively expressed "relative weight" (in rank order format). Multiple correlation coefficients ranged between .79 ( $p < .005$ ,  $df = 10, 29$ ) and .94 ( $p < .001$ ), indicating between 62% and 88% of total variance in the dependent (judgment) variable was accounted for by a linear combination of the 10 traits.

Spearman's rank order correlation ( $r_s$ ) indexed the extent of agreement between subjects' empirically *derived* beta weights (disregarding signs) and their corresponding subjectively *expressed* (introspective) weights. Values of  $r_s$  ranged between .84 ( $p < .005$ ,  $df = 8$ ) and  $-.13$  (n.s.), with a mean value of .55. Eight of the 14 comparisons were significant at  $p < .05$ .

In order to provide an overall picture of cue utilization for the class as a whole, two composite models were created. The regression model for all subjects was  $.21X_1 + .26X_2 + .09X_3 + .38X_4 + .17X_5 + .38X_6 + .26X_7 + .25X_8 + .55X_9 + .42X_{10}$ , where  $X_1 \dots X_{10}$  refers to the 10 traits previously given in Table 1. Multiple correlation was .958 ( $p < .001$ ,  $df = 10, 29$ ), again suggesting a linear cue utilization judgment pattern.

A composite model for the subjectively expressed weighting scheme was obtained by averaging the rank order of traits for all subjects. The mean values for traits  $X_1 \dots X_{10}$  were 8.14, 5.36, 6.21, 3.00, 7.29, 5.21, 6.29, 4.93, 3.29, and 5.29. Spearman's rank order correlation ( $r_s$ ) yielded an index of agreement of .673 ( $p < .25$ ,  $df = 8$ ) between the empirical and subjective schemes.

### D. DISCUSSION AND CONCLUSION

The empirically derived regression models clearly indicate different cue utilization patterns for all subjects. It would appear that implicitly held definitions of an "effective instructor" differ in terms of saliency (i.e., relative weight) for a given set of descriptive traits. Table 3 shows the number of times each trait was rated as either "most important" (first or second) or "least important" (ninth or 10th) in both empirical and subjective models.

TABLE 3  
FREQUENCY OF TRAIT RATINGS AS MOST IMPORTANT OR LEAST IMPORTANT  
FOR EMPIRICAL AND SUBJECTIVE SCHEMES

Trait number	Empirical scheme		Subjective scheme	
	Most	Least	Most	Least
1	0	4	0	8
2	1	1	1	1
3	0	3	0	2
4	4	4	9	0
5	1	7	0	4
6	4	2	3	4
7	2	1	2	4
8	2	4	3	0
9	8	2	7	0
10	6	2	2	3

While only two students *said* "grading" was most important in evaluating instructor effectiveness, six students *actually* placed major weight on the trait. On the other hand, closer agreement is found with such traits as "effectiveness of speech," "treatment of questions asked in class," and "ability to stimulate classroom discussion." Of course, no attempt should be made to generalize specific weights beyond the context of this study without additional qualification.

Because individual regression equations had moderately high multiple correlation coefficients, it seems that some evidence exists for assuming that subjects were, in fact, attending to the *same* set of instructor traits throughout the judgment task. It is not possible, however, to determine the extent to which traits other than those presented were relied upon in making the evaluations. The evidence suggests that subjects did maintain a relatively consistent judgment scheme, even though individual differences are apparent.

While subjects were moderately successful in subjectively expressing their cue utilization patterns, some exceptions appear. Specifically, recalling Table 1, trait #4 ("ability to present material on understandable plane") was *said* to be the most important consideration, while trait #1 ("apparent knowledge of subject" actually had the highest regression coefficient. Similarly, trait #10 ("grading") had the second highest regression weight, yet was only *said* to be fifth in importance, subjectively speaking. The greatest discrepancy between the empirical and subjective composite schemes was trait #8 ("interest and enthusiasm in subject"). While care must be exercised in drawing inferences from these results, the findings clearly suggest additional vigilance to subjectively expressed models of student-faculty evaluations.

When evaluations are based on simply averaging responses across a series of items, some system for weighting each item seems desirable. It may be necessary, in fact, to determine such weights separately for every teaching department or possibly every course. It would also seem worthwhile to establish some baseline or normative reference model as a practical guide. While no amount of empirical sophistication can overcome certain perturbations in ordinary student-faculty-course evaluations, the additional attention to individual differences in cue utilization seems likely to enhance the validity—and decision-making value—of accountability programs.

Regardless of the particular theoretical or methodological paradigm adopted, it seems safe to assert that educational evaluation is inevitable. As Miller (12) concluded, "Evaluation of some sort, by someone, does take place. The question is whether the procedures used and the individuals using them constitute an optimal process" (p. 6).

#### REFERENCES

1. AUSTIN, A. W., & LEE, C. B. T. Current practices in the evaluation and training of college teachers. *Educ. Rec.*, 1966, **47**, 361-365.
2. BOGUE, E. G. Student appraisal of teaching effectiveness in higher education: Summary of the literature. *Educ. Quest*, 1967, **11**, 6-10.
3. BORICH, G., Ed. Accountability. *J. Res. & Devel. Educ.*, 1971, **5**, 1-96.
4. DRESSEL, P. L., et al. *Evaluation in Higher Education*. Boston: Houghton Mifflin, 1961.
5. GULLIKSEN, H. The structure of individual differences in optimality judgments. In Shelly, M. W., & Bryan, G. L., *Human Judgments and Optimality*. New York: Wiley, 1962. Chap. 5.
6. GUSTAD, J. W. Evaluation of teaching performance: Issues and possibilities. In Lee, C. B. T. (Ed.), *Improving College Teaching*. Washington, D.C.: Amer. Council Educ., 1967. Pp. 265-281.
7. HARVEY, J. N., & BARKER, D. G. Student evaluation of teaching effectiveness. *Improv. Coll. & Univ. Teaching*, 1970, **18**, 275-278.
8. HOFFMAN, P. J. The paramorphic representation of clinical judgment. *Psychol. Bull.*, 1960, **57**, 116-131.
9. HOLMES, D. S. The teaching assessment blank: A form for the student assessment of college instructors. *J. Exper. Educ.*, 1971, **39**, 34-38.
10. HURSCH, C. J., HAMMOND, K. R., & HURSCH, J. L. Some methodological considerations in multiple-cue probability studies. *Psychol. Rev.*, 1964, **71**, 42-60.
11. MEREDITH, G. M. Dimensions of faculty-course evaluation. *J. of Psychol.*, 1969, **73**, 27-32.
12. MILLER, R. I. *Evaluating Faculty Performance*. San Francisco: Jossey-Bass, 1972.

*College of Communications*

*University of Illinois at Urbana-Champaign*

*103 Gregory Hall*

*Urbana, Illinois 61801*



# A CHILD-ADULT RESEARCH FORM OF THE PITTSBURGH SCALES OF SOCIAL EXTRAVERSION-INTROVERSION AND EMOTIONALITY\*

*Vanderbilt University*

---

NICHOLAS A. SIEVEKING<sup>1</sup>

---

## SUMMARY

Bendig's (1) Pittsburgh Scales of Social Extraversion-Introversion (SEI) and Emotionality (Em) were translated so that they were of appropriate reading difficulty for children as well as for adults. The revised scales correlated highly (both .90) with the Pittsburgh scales, and they had high test-retest reliabilities (.92 and .89). Data from samples representing third grade to college ages were presented for the revised scales. SEI-Em correlations were nonsignificant for most subsamples, indicating that the revised SEI-Em scales are independent. Scale means were stable across ages. The revised scales should be considered as research instruments, since more representative normative data and validity data are still required. The scales should, though, facilitate developmental research in extraversion and emotionality.

---

## A. INTRODUCTION

Bendig (1) developed the Pittsburgh Scales of Social Extraversion-Introversion (SEI) and Emotionality (Em) by administering items from previous inventories to American college students. The SEI and Em factors were reliable ( $K-R\ 20\ rs = .88$ ) and were relatively independent, as would be predicted from Eysenck's (4) personality theory. Bendig stated that his scales provided factorially validated, shorter, and more reliable research measures with a broader item content than were present in the scales from which he retained items (Maudsley Personality Inventory, Guilford Zimmerman Temperament Survey, Minnesota Multiphasic Personality Inventory).

Bendig's scales are attractive psychometrically, and the two personality factors that they measure relate to many research questions. The items on

---

\* Received in the Editorial Office on October 10, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Reprints and questionnaire items can be obtained from the author at the address shown at the end of the article.

these scales, however, contain such difficult words and concepts that a number of them appear appropriate only for subjects with strong vocabularies. It seemed desirable, therefore, to translate the scales downward in reading difficulty while preserving item meanings so that they could be used with younger subjects as well as with adults.

## B. PROCEDURE

The school grade levels of 340 words (nouns, verbs, adjectives, and adverbs) on the 60 Pittsburgh items were determined by averaging their placements on the Buckingham and Dolch (2) list. That list reflects the grades at which children used words spontaneously in speech and writing. An attempt was made then to translate all words downward (e.g., "meditate," grade 8, became "think," grade 2). The Buckingham-Dolch list appeared to be the most adequate for this purpose, yet because of its age all words on the rewritten items were checked against a more recent list (9) which noted the grades at which words were introduced into children's reading materials.

Three classes of introductory psychology students at Vanderbilt University received two administrations of the scales with a 96-hour interval between administrations. One class took the original scales twice, and another class took the revised scales twice. A third class took the original and the revised scales, with order of administration reversed for one-half of the sample. The revised scales were also administered to 523 students in rural and small town schools of Lawrence County, Tennessee, grades 3 through 8, and to a sample of 292 fifth graders in three Nashville, Tennessee, public schools. In grades 3 and 4, the teacher read the items to the students, who then marked their answers.

In order to have their characteristics evaluated in an atypical population, the revised scales were administered to 238 high school students at Tennessee Preparatory School (TPS), a residential school for children who are court-committed, most often for being homeless and neglected.

The 30 revised items for both the SEI and the Em scales were scored as were Bendig's original items. High scores reflect extraversion and high emotionality. Inspection of the SEI-Em scale correlations for both sexes in all samples showed that many were lower than in Bendig's sample ( $r = -.14$ ), but that several were comparatively higher ( $r$ s in the  $-.30$ 's). A new scoring system was developed by examination of item correlations with the total scale scores in two samples, those of Vanderbilt students and of Nashville fifth graders. Items that were contributing substantially to the negative SEI-Em correlations in these two samples (five SEI and two Em items) were eliminated. The revised scoring system resulted in decreased SEI-Em correlations in all

four samples. All data presented on the revised scales are from the modified scoring system.

### C. RESULTS

As is shown in Table 1, the wording revisions resulted in an appreciable decrease in difficulty. It is unlikely, however, that these statistics sufficiently reflect the decreased difficulty level, since rewordings of difficult concepts did not always result in the use of lower grade words (e.g., "crossing my bridges before I come to them," average word grade level 1.9, became "worrying about things before they happen," average grade 3.6).

Test-retest with original and revised scales in the Vanderbilt samples showed that the original SEI and Em scale reliabilities were both .94; the revised SEI and Em scales had reliabilities of .92 and .89, respectively. The original SEI and Em scales both correlated .90 with the corresponding revised scales (the original-revised correlations were .97 and .98 after corrections for unreliability). Both forms were, therefore, reliable scales measuring similar things, and they may be termed equivalent forms of the same test, differing only in wording and thus in reading difficulty.

Table 2 presents scale characteristics for original and revised forms on first and second Vanderbilt administrations. All SEI-Em correlations were negative, although none were significant ( $ps > .05$ ) for the revised form. There were no significant sex differences in correlations. Means of the revised scales are slightly lower than those for the Pittsburgh scales because of the deletion of items, noted above. The alpha, or internal consistency, coefficients for both scales of both forms varied from .84 to .91.

Analyses of variance for the Lawrence County sample (Table 3) showed that there were neither sex, grade, nor sex by grade differences in mean revised SEI or Em scores (all  $F_s < 1$ ). There were no significant SEI-Em correlations and no sex differences in correlations. The correlation over all six grades between SEI and Em scales was .00 ( $N = 523$ ). These findings suggest that the scales were measuring variables that were both largely independent of age and of one another in the age range sampled in Lawrence County. Stability of scale results across samples is suggested by similar and not significantly different ( $ps > .05$ ) SEI and Em means in the fifth graders sampled in Lawrence County and in Nashville (Table 4). In the Nashville sample, however, females had a significant ( $p < .01$ ) SEI-Em correlation which was significantly ( $p < .01$ ) different from the male SEI-Em correlation. The SEI and Em alpha coefficients for Nashville fifth graders were .78 and .80.

In the atypical TPS sample (Table 5), the sexes had nearly equal mean SEI scores, but females had significantly ( $t = 4.82$ ,  $p < .01$ ) higher Em scores. There was also a significant SEI-Em correlation for females and a signi-

TABLE 1  
 CUMULATIVE PERCENTAGES ACROSS GRADES OF WORDS ON ORIGINAL AND REVISED SCALES,  
 DERIVED FROM BUCKINGHAM-DOLCH (B-D) AND TAYLOR-FRACKENPOHL (T-F) WORD LISTS

Scale	Word list	Mean grade	SD	p <sup>a</sup>	Percentage of words at or below grades									
					1	2	3	4	5	6	7	8	9	10+ <sup>b</sup>
Original	B-D	3.6	2.1	0	30	56	65	72	81	90	97	98	99	100
Original	T-F	3.0	2.4	16	34	50	67	74	81	89	94	98	99	100
Revised	B-D	2.6	1.1	0	35	73	86	96	100	—	—	—	—	—
Revised	T-F	1.9	1.5	17	46	66	88	96	98	99	100	—	—	—

<sup>a</sup> Preprimer.

<sup>b</sup> Difficult unlisted words (e.g., "introspective") were entered at 10th grade level even though they might have been more advanced.



TABLE 2

SOCIAL EXTRAVERSION-INTROVERSION (SEI) AND EMOTIONALITY (Em) CHARACTERISTICS FOR ORIGINAL AND REVISED FORMS—FIRST AND SECOND ADMINISTRATIONS, VANDERBILT SAMPLE

Administration	SEI		Em		$r_{\text{SEI-Em}}$
	Mean	SD	Mean	SD	
Original					
Males					
First	16.8	5.8	15.2	6.9	— .16
Second	19.1	6.5	14.5	7.5	— .39*
Females					
First	19.4	6.3	14.8	6.0	— .15
Second	20.2	7.2	15.0	6.4	— .35*
Total					
First	18.1	6.2	15.0	6.4	— .16
Second	19.6	6.9	14.7	7.0	— .36*
Revised					
Males					
First	16.4	5.5	16.0	6.2	— .11
Second	15.6	5.9	15.3	7.4	— .09
Females					
First	16.8	5.0	16.7	5.2	— .20
Second	17.4	3.4	15.6	6.3	— .01
Total					
First	16.5	5.3	16.3	5.9	— .14
Second	16.3	5.1	15.4	6.9	— .06

Note: *N*s were as follows—Original form, males 55, females 51; Revised form, males 48, females 32.

\*  $p < .01$ .

ficant ( $p < .01$ ) sex difference in SEI-Em correlation. Additionally, Em was significantly negatively correlated with MMPI-L scores for males. The SEI and Em alpha coefficients for the TPS sample were .83 and .76.

Two way analysis of variance across all samples, with classifications by sex and by school year (with the exception of the Vanderbilt sample in which all four years were combined), were performed for the SEI and the Em scores. There were no significant ( $p > .05$ ) main effects of sex on SEI scores (means = 17.0, male; 17.2, female) and no significant interaction of sex with school year on either SEI or Em scores. The main effects of sex on Em scores were significant ( $p < .001$ ;  $F = 46.6$ ;  $df = 2$  and 1118) and showed that females overall obtained higher Em scores than did males (means = 18.4 and 16.4, respectively). There were significant main effects of school year on both SEI ( $p < .05$ ;  $F = 1.94$ ;  $df = 11$  and 1118) and Em ( $p < .001$ ;  $F = 5.96$ ;  $df = 11$  and 1118). There were, however, no individual mean differences within the main effects of school year for either SEI or Em as tested with Scheffé's method at alpha = .01. When alpha was set at .05, there were again

TABLE 3  
REVISED SOCIAL EXTRAVERSION-INTROVERSION (SEI) AND EMOTIONALITY (Em) DATA  
FROM GRADES 3 TO 8, LAWRENCE COUNTY, TENNESSEE

Grade and Sex	N	SEI		Em		$r_{SEI-Em}$
		Mean	SD	Mean	SD	
3						
Males	47	17.8	3.4	14.4	5.1	-.14
Females	47	17.0	4.2	16.5	5.0	-.01
Total	94	17.4	3.9	15.4	5.2	-.09
4						
Males	44	17.8	3.8	17.8	4.6	.11
Females	38	17.4	3.8	18.9	4.4	-.14
Total	82	17.6	3.8	18.3	4.5	-.01
5						
Males	32	17.5	3.6	15.4	4.1	.03
Females	41	17.2	3.6	17.8	4.9	-.05
Total	73	17.4	3.6	16.8	4.7	-.03
6						
Males	49	16.6	4.8	15.1	5.8	.20
Females	50	17.1	4.3	17.0	5.3	-.08
Total	99	16.9	4.6	16.0	5.6	.08
7						
Males	44	16.4	4.2	16.5	4.5	.05
Females	47	17.3	4.5	19.5	4.8	-.14
Total	91	16.8	4.3	18.1	4.8	-.02
8						
Males	43	17.3	4.2	15.2	3.4	.06
Females	41	18.2	4.3	18.9	4.8	-.06
Total	84	17.8	4.3	17.0	4.5	.03
Total 3-8						
Males	259	17.2	4.1	15.7	4.8	.07
Females	264	17.4	4.2	18.0	5.0	-.07
Total	523	17.3	4.1	16.9	5.1	.00

no differences among SEI means, but there were differences among Em means. TPS eleventh graders ( $\bar{X} = 19.9$ ) differed from Vanderbilt students ( $\bar{X} = 15.6$ ) and from Lawrence County third graders ( $\bar{X} = 15.6$ ); and Lawrence County sixth graders ( $\bar{X} = 18.7$ ) differed from Vanderbilt students.

Correlations with scales of another self-report measure, the Tennessee Self Concept Scale (6) were obtained on a subsample of 50 TPS students. One-half of these students were above the 70th percentile (extraverts) and one-half were below the 30th percentile (introverts) on SEI, and all had volunteered for treatment of interpersonal anxiety (8). SEI was related positively ( $p < .01$ ) to scales reflecting the concepts of the physical self ( $r = .40$ ), personal self ( $r = .46$ ), family self ( $r = .35$ ), and social self ( $r = .56$ ). Em was significantly related only to the personal self-concept ( $r = -.39$ ).

TABLE 4

DATA ON REVISED SOCIAL EXTRAVERSION-INTROVERSION (SEI) AND EMOTIONALITY (Em)  
SCALES FROM NASHVILLE FIFTH GRADERS

Sex	N	SEI		Em		$r_{SEI-Em}$
		Mean	SD	Mean	SD	
Males	138	17.6	4.2	16.2	4.8	.11
Females	154	17.3	4.3	17.4	5.4	-.22*
Total	292	17.5	4.3	16.8	5.1	-.07

\*  $p < .01$ .

## D. DISCUSSION

The downward translation of item wordings accompanied by demonstrating equation of the new child-adult and original adult scales seemed to be a fruitful approach to test reconstruction. It is assumed that the reduced reading difficulty of the scales made them more reliable and valid at the lower grade levels. The scales produced were stable across a wide range of age groups, and the SEI-Em scores derived were more independent than in the original scales. The approach used makes retention of separate child and adult SEI and Em scales unnecessary, since the revised scales are applicable to diverse age groups. The use of identical scales for all ages should facilitate both longitudinal and cross-sectional studies.

The generally higher Em scores from the TPS subjects provide some evidence for validity, since these subjects were believed by their teachers to be more anxious than is typical for their age groups and were generally high on other research measures of anxiety (8). However, no normal high school group was used for comparison with the TPS sample. Some evidence for the relation of SEI to other measures was found also in its correlation with various aspects of the self-concept in the atypical TPS sample. It should be noted also that both SEI and Em were very highly correlated with Bendig's original scales, the

TABLE 5

DATA ON REVISED SOCIAL EXTRAVERSION-INTROVERSION (SEI) AND EMOTIONALITY (Em)  
SCALES FROM TENNESSEE PREPARATORY SCHOOL STUDENTS

Sex	N	SEI		Em		$r_{SEI-Em}^a$
		Mean	SD	Mean	SD	
Males	128	16.7	4.9	17.7	4.6	-.06
Females	110	16.3	5.2	20.4	4.1	-.36*
Total	238	16.4	5.1	18.8	4.6	-.19*

<sup>a</sup>  $r$ s with MMPI-L scale (mean = 3.2; SD = 2.0): Males SEI-L -.11, Em-L -.32\*; Females SEI-L -.09, Em-L -.18; Total SEI-L -.11, Em-L -.23\*.

\*  $p < .01$ .

items of which were drawn from better established tests. Evidence for validity is, however, very incomplete. Future attention must be paid to validation of both scales, especially to their correlation with peer ratings and with behavioral observations.

There is a history of disputes regarding the independence of extraversion-introversion and neuroticism (emotionality). In his review of the relevant literature, Carrigan (3) concluded that those two factors were not unequivocally proven independent, although other authors do report essential independence (e.g., 5). Such independence is important to Eysenck's (4) theory of personality, and for other reasons it is desirable psychometrically. In the present study it was possible to derive SEI and Em scales that were only slightly related for most of the subsamples. The highest correlation was found in the atypical TPS sample. Jensen (7) noted also that significant correlations between scales similar to SEI and Em were found in some abnormal groups.

The lowest age limit of applicability of the scales needs to be determined by reading the items to subjects below third grade level. Norms and factor structure across all ages should be derived. Wider geographical sampling is required. The revised scales are, however, promising for research use. They seem applicable to broad age ranges and should help to answer questions about developmental changes in extraversion-introversion and emotionality.

#### REFERENCES

1. BENDIG, A. W. The Pittsburgh Scales of Social Extraversion-Introversion and Emotionality. *J. of Psychol.*, 1962, **53**, 199-209.
2. BUCKINGHAM, B. R., & DOLCH, E. W. A Combined Word List. Boston: Ginn, 1936.
3. CARRIGAN, P. M. Extraversion-introversion as a dimension of personality: A reappraisal. *Psychol. Bull.*, 1960, **57**, 329-360.
4. EYSENCK, H. J. Dimensions of Personality. London: Routledge & Kegan Paul, 1960.
5. FARLEY, F. H. Moderating effects of intelligence on the independence of extraversion and neuroticism. *J. Consult. & Clin. Psychol.*, 1968, **32**, 226-228.
6. FITTS, W. H. Tennessee Self Concept Scale. Nashville, Tenn.: Counsel. Recordings & Tests, 1965.
7. JENSEN, A. R. The Maudsley Personality Inventory. *Acta Psychol.*, 1958, **14**, 314-325.
8. SIEVEKING, N. A. Group systematic desensitization with institutionalized adolescents. Unpublished Doctoral dissertation, University of Illinois, Urbana, 1969.
9. TAYLOR, S. E., & FRACKENPOHL, H. A core vocabulary. EDL Research and Information Bulletin No. 5, Educational Developmental Laboratories, Huntington, New York, 1960.

*Interuniversity Psychological and Counseling Center*  
*Vanderbilt University*  
 300 Oxford House, 1313-21st Avenue South  
 Nashville, Tennessee 37212



## BEHAVIORAL VARIABILITY AMONG RETARDATES, CHILDREN, AND COLLEGE STUDENTS\*

*University of Alabama*

---

ANN M. LIEBERT AND ALFRED A. BAUMEISTER

---

### SUMMARY

Individuals seem to differ in their ability to maintain consistent performance. This is obviously an important attribute of behavior that has been largely neglected by researchers. There is, however, some evidence to suggest that, in certain types of situations, the behaviors of the mental retardate are characterized by a relatively high degree of variability. Two studies are reported here that deal with some of the developmental implications of intraindividual variability. In one study, involving retardates, reliability and generality of response consistency were examined in reaction time, time estimation, and digit span tasks. It was shown that individual differences in consistency are reliable and that these generalize across tasks. A second experiment was concerned with the developmental aspects of variability. First-, third-, and fifth-grade pupils and college students were given many trials on a reaction-time task. A variety of measures, reflecting typical performance, limit of performance, and variability, were computed for each subject. The ability to react quickly shows a developmental trend not only with respect to overall performance, but also in the efficiency with which the individual maintains his optimal level of responding.

---

### A. INTRODUCTION

A focal point for theoretical considerations and empirical investigations of retardate performance based on intraindividual variability has been proposed by Baumeister (1). This approach, with its emphasis on variability, differs from most behavioral characterizations of mental retardation which are derived from a level-of-performance conceptualization. With respect to the latter, the retardate is characterized as slower, poorer, weaker, or inferior in performance than the normal individual. However, analyses of intraindividual variability

---

\* Received in the Editorial Office on October 13, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

offer an additional aspect, consistency of performance, for our conceptualization of retarded behaviors.

Variance, as an aspect of behavior development and inadequacy, has received relatively little attention. While it is often noted that retarded groups are more heterogeneous in their performance than normal groups, this variability has not been systematically scrutinized. Usually it is attributed to individual or between-subject differences. That such differences can be observed is not really surprising, in that retarded behavior probably has its origins in a wide array of processes. Yet, there is another way of thinking about variability—that variability associated with an *individual's* inconsistency in performance from moment-to-moment, from trial-to-trial. This component of variance, intraindividual variability, is assumed to be associated with changing internal processes of the individual. Specifically, intraindividual variability refers to alterations in performance observed at different points in time when objective conditions are held constant and no directional response trends are evident.

While variability in behavior is obviously adaptive for psychological development, in a particular task situation excessive variability is not conducive to good overall performance. Adaptation of the individual to his environment is enhanced through certain acts which take the same form repeatedly as they are more rapidly and efficiently executed. Where the quality of a performance is specified by a high performance score, good overall performance is characterized by high level with minimal variability. But, obviously consistency alone is not adequate for good performance, inasmuch as minimal variability may accompany poor level. In other words, some individuals may demonstrate consistently poor performances.

Variability and level are associated aspects of performance, though not in a precise one-to-one relationship. Two factors then are implicated in evaluation of an individual's performance: (a) limit of performance and (b) efficiency in responding at this limit. Efficiency, the construct invoked by Berkson and Baumeister (4), denotes an individual's ability to maintain his performance at his optimal level. Maximal efficiency therefore refers to minimal variability from one's limit. Inferences about behavior are usually made in terms of limit, but the position taken here is that efficiency may have particular relevance for theoretical conceptualizations of behavior development in general and retardation in particular.

Though the importance of intraindividual variability has been recognized by numerous theorists (5, 6, 9, 11, 14), certain commonly employed statistical methods conceal and preclude systematic investigation of intraindividual variability. It is general practice to pool numerous scores of an individual to obtain

a single representative estimate of his performance. These scores, in turn, are combined with the scores of other individuals, and the group average obtained. Thus information relating to intersubject variance may be obtained (as a conventional error term), but the contribution of intraindividual variability to the average score of the individual is obscured. It is clear, however, that measurement of intraindividual variability may provide cues pertaining to the effectiveness of a person's overall performance.

The general findings of recent investigations of intraindividual variability suggest that efficiency is a reliable and predictable characteristic of the individual (2, 4). Large differences between intelligence groups in relation to efficiency have been found, indicating that retardates exhibit greater variability in performance in comparison to normal subjects. Low efficiency and depressed level both appear necessary to describe retardate performance (3, 8) on many tasks. Moreover, manipulation of subject and task variables systematically influences efficiency (2, 12).

While low efficiency has been shown to be associated with retardate functioning on certain tasks—e.g., reaction time (2, 3, 4, 12), auditory discrimination (8), and short-term memory (2)—no systematic study of the generality and reliability of intraindividual variability has been reported. Furthermore, though significant quantitative differences between the efficiency of normal and retarded groups have been found, an investigation of the developmental aspects of intraindividual variability has yet to be done. The following experiments provide data pertaining to the generality, reliability, and developmental aspects of intraindividual variability.

## B. STUDY I

This experiment was designed to assess the contribution of intraindividual variability to the typical performance of retardates and, further, to evaluate the reliability and generality of intraindividual variability as an appropriate characterization of retardate behavior. Three tasks, selected to measure differing processes (short-term memory, temporal judgment, and reaction time) and amenable to multiple measurement without marked sequential effects, were employed.

### 1. *Method and Subjects*

Thirty retardates (20 male, 10 female) were selected from the population of a state residential institution for the retarded. None showed any gross sensory or motor defects. The median age of the group was 18.1 years with a range of 13.0 to 56.4 years. Data secured from institutional records indicated

the group to have a median *MA* of 8.9 years with a range of 4.9 to 12.7 years. The same *Ss* were administered all three tasks twice, the second testing within four days of the first.

## 2. Task I: Reaction Time (RT)

*a. Apparatus.* The apparatus consisted of a Beltone audiometer and headset, a recycling Hunter Timer, a telegraph key for responding, and a Hunter Klockounter.

*b. Procedure.* The *S* was seated before a table upon which the telegraph key was mounted. The *E* determined the *S*'s threshold for a 1000 cps pure tone and set the audiometer at 50 db above threshold. The *S* was then instructed to release the depressed key upon hearing the signal ton in his ear. Speed of response was emphasized. The *E* demonstrated the task for each *S* and provided five practice trials to assure comprehension of the task. Fifty trials were run with a five-second interval between trials, during which the *S* depressed the key in preparation for the next tone. Reaction time was recorded to the nearest .001 second.

## 3. Task II: Digit Span (DS)

*a. Materials.* The material for this task consisted of a series of 12 lists of random numbers. Each list began with a two-digit number, with each series increasing in length by one digit until the final nine-digit number was reached. No numbers were repeated within a series.

*b. Procedure.* The *S* was seated before the *E* and instructed to listen to the numbers which *E* read and to repeat the digits when *E* had finished. One or two practice trials were conducted as required for the particular *S*. The *E* read the digits at the rate of one per second. When *S* failed to repeat the digits correctly, *E* recorded the last number of digits correctly reproduced and proceeded to the next list until 12 trials had been given.

## 4. Task III: Time Estimation (TE)

*a. Apparatus.* Programming equipment for the time estimation task included a Hunter Interval Timer calibrated for intervals of 1, 3, 5, 7, and 9 seconds duration, a one-inch circular red light (120 volt, 15 watt) within a dark 1.5 × 2.0-inch rectangle on the display unit, and a Hunter Klockounter.

*b. Procedure.* The *S* sat at a table two feet from the mounted display unit. He was instructed to observe the stimulus light and note how long it appeared. Upon its termination, *S* depressed the response key located directly in front of him. This key turned on the stimulus light again, and *S* was to keep it on for the same amount of time that the original stimulus had appeared. Each *S*



practiced the task for five trials. Each of the five time intervals appeared eight times in a random order for a total of 40 trials. The same presentation order was used for every *S*. Between each trial there was a brief interval during which *E* recorded *S*'s response to the nearest .01 second and adjusted the timer for the next estimation duration.

### 5. Results

Mean, median, and standard deviation scores were calculated for each *S* on each of the tasks. As a particular task was administered on two occasions, there were two sets of scores available for each *S*. To ascertain the reliability of the measurements across time, the first test administration scores of each *S* were correlated with those of the second administration. Table 1 presents test-retest reliability coefficients with respect to means, medians, and standard deviations of each of the three tasks. With the one exception of DS *SD* ( $r = .13$ ) all correlations were significant ( $p < .01$  to  $< .001$ ).

TABLE 1  
TEST-RETEST RELIABILITY

Task	$\bar{X}$	<i>Mdn</i>	<i>SD</i>
RT	.70**	.74**	.55*
TE	.71**	.70**	.49*
DS	.71**	.72**	.13

\*  $p < .01$ .

\*\*  $p < .001$ .

The correlations of measurements within specific tasks were calculated as a test of the contribution of efficiency to typical performance. The correlations of major interest are those between the measures of central tendency and *SD* on each task. These results along with the intertask correlations are presented in Table 2. On the RT and TE tasks, correlations of central tendency and variability measures were significant ( $p < .02$  to  $< .01$ ). The correlations of standard deviations between the different tasks were examined as a test of the generality of the concept of efficiency in relation to retardate performance. One significant correlation was apparent: RT *SD* and TE *SD*,  $r = .40$  ( $p < .05$ ).

### C. STUDY II

The second study was designed to investigate the development of efficiency in normal school children. Qualitative and quantitative developmental changes in children's reaction times are apparent, but explanations in terms of improvement due to differences in motor coordination or motivation do not seem to

TABLE 2  
INTRA- AND INTERTASK CORRELATIONS

Measures	RT $\bar{X}$	RT <i>Mdn</i>	RT <i>SD</i>	TE $\bar{X}$	TE <i>Mdn</i>	TE <i>SD</i>	DS $\bar{X}$	DS <i>Mdn</i>	DS <i>SD</i>
RT $\bar{X}$		.98***	.64***	.29	.25	.33	-.20	-.20	-.12
RT <i>Mdn</i>			.54***	.28	.24	.32	-.22	-.20	-.15
RT <i>SD</i>				.25	.20	.40*	-.06	-.10	.00
TE <i>Mdn</i>					.66***	.52***	-.12	-.12	.09
TE $\bar{X}$						.44**	-.08	-.13	.13
TE <i>SD</i>							-.22	-.27	-.10
DS $\bar{X}$								.95***	.34
DS <i>Mdn</i>									.34
DS <i>SD</i>									

\*  $p < .05$ .

\*\*  $p < .02$ .

\*\*\*  $p < .01$ .

account for the marked decreases in reaction time with age (10). Data concerning efficiency in normal children may suggest possible explanations of the developmental changes in reaction time among children.

### 1. Method

*a. Subjects.* Twenty-seven children from an upper-middle-class elementary school (nine in each of the first, third, and fifth grades) and nine volunteer college students served as Ss in the study.

*b. Apparatus.* The apparatus was the same as that used in RT task of Study I.

*c. Procedure.* The procedure was similar to that of the RT task of Study I. However, each subject was tested on three different days with 50 trials each day.

### 2. Results

The RTs of the first day were not employed in the analysis, but were regarded as practice trials. The mean on each S's 10 fastest scores was arbitrarily selected as that S's lower limit of performance (LL). Each S's remaining 90 reaction time responses were included in calculations of the S's  $\bar{X}$ , *Mdn*, and *SD* scores. Analyses of variance indicated significant age trends for all measures of RT: LL, deviation from LL (LLD),  $\bar{X}$  *Mdn*, and *SD*. As age increased, RT improved, and intraindividual variability decreased. The means and standard deviations of the measures for the retardates and normal subjects are presented in Table 3.

Analyses of variance computed on the measures of RT across groups

TABLE 3  
MEANS AND STANDARD DEVIATIONS OF REACTION TIME MEASURES FOR  
NORMALS AND RETARDATES (STUDIES I AND II)

Measure	Retardates	First grade	Third grade	Fifth grade	College students
Mean					
$\bar{X}$	.475	.436	.371	.286	.199
$SD$	.150	.046	.078	.056	.083
Median					
$\bar{X}$	.455	.419	.358	.276	.194
$SD$	.155	.044	.074	.052	.080
Standard deviation					
$\bar{X}$	.135	.128	.100	.071	.039
$SD$	.065	.027	.046	.020	.018
Lower limit (LL)					
$\bar{X}$	.277	.230	.214	.181	.144
$SD$	.070	.080	.026	.044	.061
Lower limit deviation (LLD)					
$\bar{X}$	.223	.229	.174	.116	.062
$SD$	.105	.073	.092	.027	.031

revealed significant differences between the age groups on all five measures ( $p < .05$  to  $p < .005$ ). These results are presented in Table 4. Additional  $t$ -tests comparing each group's performance with every other group indicated that the  $\bar{X}$ ,  $Mdn$ , and  $SD$  of each group differed significantly from every other group. However, the significant  $F$  for the LL can be attributed to the differences in LL between the college students and the first- and third-grade children. In reference to the LLD, the college students and fifth-grade children were significantly less variable than the first and third graders. Correlations between  $\bar{X}$  and  $SD$  scores for the first-, third-, fifth-grade, and college students were, .34, .96, .85, and .91 for the four groups, respectively. Correlations between  $\bar{X}$  and LL were .58, .01, .91, and .96 for the four groups, respectively.

#### D. DISCUSSION

The significant intratask correlations of the retardates' average performance and variability in the first study are in agreement with the findings of earlier studies, and support the conceptualization of retardate performance in terms of efficiency. These positive correlations suggest that low performance may in part be due to the individual's inability to perform efficiently. The absence of high test-retest reliability for the  $SD$  of the DS task, together with the high reliability of the  $\bar{X}$  and  $Mdn$  scores for that task, suggests that DS is essentially

TABLE 4  
SUMMARY ANALYSES OF VARIANCE

Source	<i>df</i>	<i>MS</i>	<i>F</i>
Mean			
Between	3	.074	22.893**
Within	32	.003	
Total	35		
Median			
Between	3	.067	23.232**
Within	32	.003	
Total	35		
Standard deviation			
Between	3	.012	13.898**
Within	32	.001	
Total	35		
Lower limit			
Between	3	.008	3.113*
Within	32	.002	
Total	35		
Deviation from lower limit			
Between	3	.043	11.511**
Within	32	.004	
Total	35		

\*  $p < .05$ .\*\*  $p < .005$ .

a measure of limit. On this task there was little fluctuation in the *SD* of performance within subjects, indicating that these individuals were performing consistently at their limits.

Of equal importance was the significant intertask correlation of *RT* and *TE SDs*. This correlation specifically points to the generality of intraindividual variability, suggesting that an individual who is variable in one context may be variable in another. Obviously, a process common to both tasks may be the ability to respond on the basis of temporal cues. This may be a process particularly sensitive to individual differences in efficiency.

The results of Study II showed a definite increase in efficiency with age along with overall improvement in reaction time ( $\bar{X}$ , *Mdn*, *SD*, *LL*, and *LLD* decrease with age). The average reaction time is a function of both a hypothetical limit of speed and the intraindividual variability of scores in a series of trials. Thus, developmental changes in reaction time should be explained in terms of both of these factors. The relation between variability and average performance is evident in the significant correlations between  $\bar{X}$  and *SD* for the third- and fifth-grade pupils and college students. Berkson and Baumeister (4) explain that in the case of each slow subjects who do not approach their limit, there may be a correlation between variability and average performance.



The nonsignificant correlation between  $\bar{X}$  and  $SD$  for first graders could be explained in this manner.

Efficiency seems to follow a developmental course, but the correlates of this basic factor still remain to be identified. Since objective environmental determinants have been minimized, the response variability appears to be associated with the changing internal characteristics of the individual. Variability may be related to psychobiological aspects of the individual's functioning. Recent investigations of variability by neurophysiologists (7, 13) have provided strong supportive evidence of a relationship between variability and ongoing internal processes.

#### REFERENCES

1. BAUMEISTER, A. A. Behavioral inadequacy and variability of performance. *Amer. J. Ment. Defic.*, 1968, **73**, 477-483.
2. BAUMEISTER, A. A., & KELLAS, G. A. Intrasubject response variability in relation to intelligence. *J. Abn. Psychol.*, 1968, **73**, 421-423.
3. ———. Distribution of reaction times of retardates and normals. *Amer. J. Ment. Defic.*, 1968, **72**, 715-718.
4. BERKSON, G., & BAUMEISTER, A. A. Reaction time variability of mental defectives and normals. *Amer. J. Ment. Defic.*, 1967, **72**, 262-266.
5. CATTELL, R. B. *Personality and Motivation Structure and Measurement*. Yonkers, N. Y.: World Book, 1957.
6. DODGE, R. *Conditions and Consequences of Human Behavior*. New Haven: Yale Univ. Press, 1931.
7. DONCHIN, E., & LINDSLEY, D. B. Average evoked potentials and reaction time to visual stimuli. *Electroenceph. & Clin. Neurophysiol.*, 1966, **20**, 217-223.
8. DUGAS, J. L., & BAUMEISTER, A. A. A comparison of intrasubject variability in auditory difference limens of normals and retardates. *Amer. J. Ment. Defic.*, 1968, **73**, 500-504.
9. FISKE, D. W., & RICE, L. Intra-individual response variability. *Psychol. Bull.*, 1955, **52**, 217-250.
10. HOHLE, R. H. Component process latencies in reaction times of children and adults. In L. P. Lipsitt & C. C. Spiker (Eds.), *Advances in Child Behavior and Development* (Vol. 3). New York: Academic Press, 1967. Pp. 225-261.
11. HULL, C. L. *A Behavior system*. New Haven: Yale Univ. Press, 1952.
12. KELLAS, G. Reaction-time and response variability of normal and retarded individuals. *Amer. J. Ment. Defic.*, 1969, **74**, 409-414.
13. MORRELL, L. K., & MORRELL, F. Evoked potentials and reaction times: A study of intra-individual variability. *Electroenceph. & Clin. Neurophysiol.*, 1966, **20**, 567-575.
14. THOULESS, R. H. Test unreliability and function fluctuation. *Brit. J. Psychol.*, 1936, **26**, 325-343.

Center for Developmental and Learning Disorders  
University of Alabama  
Box 6302  
University, Alabama 35486



## IMPAIRED SEX-ROLE IDENTIFICATION IN SCHIZOPHRENIA EXPRESSED IN THE COMPREHENSION OF HUMOR STIMULI\*

*Syracuse University; West Haven Veterans Administration Hospital,  
Connecticut; and Yale University*

---

JONATHAN ECKER, JACOB LEVINE,<sup>1</sup> AND EDWARD ZIGLER

---

### SUMMARY

Forty schizophrenics (20 male and 20 female) and 20 normals (10 male and 10 female) were given three tests for sex-role identification (Role Preference, Body Parts Satisfaction, and Humor). No differences between schizophrenics and normals were found on the Role Preference and Body Parts Satisfaction Tests. On the Humor Test, the schizophrenics were less capable than the normals in comprehending cartoons depicting individuals engaged in abnormal or ambiguous sex roles. No such disability was found in the schizophrenics' comprehension of cartoons involving non-sex-role or normal sex-role behavior. The findings on the Humor Test confirmed the conclusion of McClelland and Watt (6) that schizophrenics are disturbed in their sexual identification. Inconsistencies between the findings of this and the McClelland and Watt study on the Role Preference and Body Parts Satisfaction Tests were discussed.

---

### A. INTRODUCTION

Some evidence has recently been presented suggesting that schizophrenics are disturbed in their sexual identification (1, 2, 6). Unlike their normal counterparts, male schizophrenics have been found to behave in a nonmasculine manner—i.e., passive and withdrawn—whereas female schizophrenics have been found to act in a masculine way—i.e., aggressive and noisy. However, this generalization should be qualified by some findings of McClelland and Watt (6) indicating that schizophrenics and especially female schizophrenics suffer from sexual alienation rather than sexual reversal. These investigators employed both overt and covert measures of sex-role identification. No significant differences between normals and schizophrenics were found on measures of overt sex-linked interests and attitudes. However, the more covert measures

---

\* Received in the Editorial Office on October 19, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Requests for reprints should be directed to Dr. Jacob Levine at the address shown at the end of this article.

revealed significant differences indicating either sexual reversal or sexual alienation in the schizophrenics. Among the findings of this provocative study was the discovery of sexual reversal among schizophrenics of both sexes in the sex roles they chose in a fantasied role playing test. On a test of satisfaction with body parts, McClelland and Watt also found that male schizophrenics displayed a pattern like normal females, while female schizophrenics displayed a pattern similar to normal males. However, on examination of attitudes about sex-linked body parts, the findings suggested that female schizophrenics suffer from sex-role alienation rather than simple sex-role confusion.

A methodological shortcoming of the McClelland and Watt study was that their comparisons were between schizophrenics hospitalized an average of over 10 years with normal controls who were not hospitalized at all. Thus, the sex-role alienation they discovered in the schizophrenics might be attributable to hospitalization *per se* rather than to psychopathology. McClelland and Watt entertained this possibility but dismissed it on the basis of the total pattern of their findings. They also noted a second difficulty in their study in that the schizophrenic group had less education and significantly lower *IQ* scores than the normal group. McClelland and Watt maintained that these differences "reflect the disease process rather than basic differences in social-class background" (6, p. 229). However, they presented no supporting evidence for this inference.

One goal of the present study was therefore to assess the replicability of the McClelland and Watt findings obtained with their Role Preference and Body Parts Satisfaction Tests, in comparisons between groups of equally short-term hospitalized normals and schizophrenics who were also equated with respect to education and intelligence. The second goal was to determine whether the heretofore discovered sexual alienation of schizophrenics would manifest itself in respect to the comprehension of humor stimuli involving sex-role identification.

For the purposes of this investigation, two major groups of cartoons were employed: (a) non-sex-role cartoons, and cartoons in which men and women fulfilled their normal sex roles; and (b) cartoons in which the characters were engaged in abnormal or ambiguous sex roles. Considerable work has been done indicating that schizophrenics defend themselves against threatening scenes displayed in either realistic or humorous material (3). In the case of threatening cartoons, this defense typically takes the form of the schizophrenic not allowing himself to comprehend the cartoon and thus not being forced to deal with the implication of the cartoon. Even normal sophisticated persons resort to this type of defense by using either perceptual scotomazations or dis-



tortions (4). Past work with the manner in which schizophrenics deal with cartoons with threatening content, in conjunction with the hypothesis that schizophrenics suffer from sexual alienation, thus generates the prediction that schizophrenics will differ from normals in their comprehension of cartoons depicting abnormal sexual behavior. They should not differ, however, on non-threatening cartoons; i.e., those not involving sexual behavior or those in which the characters are behaving in a gender-appropriate manner.

## B. METHOD

### 1. Subjects

The subjects in this study were 40 schizophrenics (20 male and 20 female) and 20 normals (10 male and 10 female). All subjects were white, between the ages of 20 and 50, and had been hospitalized for a brief period. All schizophrenics had received an unequivocal diagnosis of schizophrenia. The normal subjects, though hospitalized medical patients, did not suffer from severe physical illness. All subjects were matched as closely as possible in respect to age, intelligence, education, and length of hospitalization. The Ammons Full-Range Picture Vocabulary Test, given prior to the experiment proper, was used to obtain intelligence quotients. The subject characteristics are presented in Table 1.

TABLE 1  
CHARACTERISTICS OF GROUPS

Group	N	Age (in years)		Education (in years)		IQ		Hospitalization (in months)	
		$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Schizophrenics									
Male	20	28.4	7.9	12.4	2.4	104	14.0	4.1	3.0
Female	20	38.8	6.2	12.0	2.1	102	11.3	4.9	4.2
Normals									
Male	10	32.0	11.2	12.5	2.6	104	9.4	2.4	3.0
Female	10	30.6	11.1	11.0	2.7	95	13.4	1.0	0.3

### 2. Procedure

Each subject was tested individually in as nonthreatening a manner as possible. Both the psychiatric and nonpsychiatric patients were told that the tests employed in this study had nothing to do with their treatment or general hospital regimen. It was emphasized that participation in the study was purely voluntary. Every subject received the Role Preference Test first, the Body

Parts Satisfaction Test second, and the Humor Test third. In no case did any subject react as though he thought the three tests were related in any way.

### 3. *Dependent Measures*

a. *Role Preference Test.* Following the procedure of McClelland and Watt (6), each subject was asked which part he would like to play if given the following choices: old grandfather *versus* old grandmother, angel *versus* Lord, scientist *versus* fashion designer, sister *versus* brother, devil *versus* witch, secretary *versus* policeman, cow *versus* bull.

b. *Body Parts Satisfaction Test.* The subject was given 20 four-inch (.1016m)  $\times$  six-inch (.1524m) cards, each of which contained the name of a part of the body. He was asked to separate the cards into two groups, one of parts with which he was satisfied and one of parts with which he was dissatisfied. As McClelland and Watt noted, previous research indicated that eight of the parts had a predominately feminine connotation, seven had a masculine connotation, and five were nonspecific with respect to gender. Three scores were obtained: total number of cards in the satisfied pile, total number of male parts cards in the satisfied pile, and total number of female parts cards in the satisfied pile. (Unlike the McClelland and Watt procedure, subjects were not instructed to make their satisfied and dissatisfied piles contain an approximately equal number of cards. It was felt that the McClelland and Watt procedure introduced the extraneous variable of the subject's ability and motivation to comply with an instruction which forces one to be either satisfied or dissatisfied with half the body parts presented.)

c. *Humor Test.* This test consisted of 20 cartoons and employed the procedure of the Mirth Response Test of Redlich, Levine, and Sohler (7). Ten of the cartoons were either irrelevant to sex-role behavior ( $N = 5$ ) or depicted men and women engaging in appropriate sex roles ( $N = 5$ ). The remaining 10 cartoons showed men and women engaged in deviant or odd sex behavior. Of these 10, five showed men behaving in a weak, passive way; three showed women behaving in a domineering or aggressive manner; and two showed characters engaging in ambiguous sex roles.

The test consisted of three successive steps. In Step 1, the subject was shown the 20 cartoons, and, following the procedure of Redlich *et al.* (7), his mirth response to each cartoon was recorded on a scale from zero (negative response) to 5 (full laugh). In Step 2, the subject sorted the cartoons into the three categories of like, indifferent, and dislike. Each cartoon was then scored in the following way: zero for dislike, 1 for indifferent, and 2 for like. In the third step, the subject was asked to explain what was funny about each of the

cartoons. Complete comprehension was scored 2, partial comprehension was scored 1, and no comprehension was scored zero.

In view of the high reliability of the mirth response reported in earlier studies (5, 8), this measure was not subjected to a reliability check. In order to assess the reliability of the comprehension scores obtained with this particular set of cartoons, the verbatim explanations of the cartoons given by 30 subjects selected at random were scored by a second judge. The reliability of the comprehension scores for those 30 subjects was found to be .95.

### C. RESULTS

#### 1. Role Preference Test

Performance of the groups on this test is presented in Table 2. Chi square analyses of male schizophrenics and normals, female schizophrenics and normals, and total schizophrenics and normals revealed no significant differences. With the exception of the tendency for male schizophrenics to make a greater number of opposite sex choices than normal males, the findings on this test are unlike those of McClelland and Watt.

TABLE 2  
NUMBER AND PERCENTAGES OF MALE AND FEMALE NORMAL AND SCHIZOPHRENIC  
SS MAKING OPPOSITE SEX CHOICES IN ROLE PREFERENCE TEST

Group	Opposite sex choices					
	Zero to 1		2		3 or more	
	N	%	N	%	N	%
Male						
Schizophrenic	13	65	3	15	4	20
Normal	7	70	3	30	0	0
Female						
Schizophrenic	17	85	3	15	0	0
Normal	7	70	2	20	1	10

#### 2. Body Parts Satisfaction Test

Performance of the groups on this test is presented in Table 3. Chi square analyses failed to reveal any significant difference between schizophrenics and normals associated with the data reported in this table.

#### 3. Humor Test

The comprehension scores of the groups for the normal (non-sex role and appropriate sex role) and abnormal sex-role cartoons are presented in Table 4. A  $2 \times 2 \times 2$  (Diagnosis  $\times$  Sex  $\times$  Type of Cartoon) repeated measures

TABLE 3  
PERCENTAGE OF MALE AND FEMALE NORMALS AND SCHIZOPHRENICS  
EXPRESSING SATISFACTION WITH PARTS OF THE BODY

Body parts	Females		Males	
	Schizophrenic	Normal	Schizophrenic	Normal
All parts ( $N = 20$ )				
Satisfied with 14 or more	70	60	85	90
Female parts ( $N = 8$ )				
Satisfied with 7 or more	40	40	65	70
Male parts ( $N = 7$ )				
Satisfied with 6 or more	70	50	85	70

analysis of variance conducted on these scores revealed significant main effects for Diagnosis ( $F = 10.10$ ;  $df = 1/56$ ;  $p < .005$ ) and Sex ( $F = 5.22$ ;  $df = 1/56$ ;  $p < .05$ ), and a significant interaction between Diagnosis and Sex ( $F = 4.11$ ;  $df = 1/56$ ;  $p < .05$ ). As can be seen in Table 4, across all cartoons normals evidenced higher comprehension scores than schizophrenics, and males evidenced higher comprehension scores than females. The Diagnosis by Sex interaction primarily reflected the fact that the comprehension scores of schizophrenic males differed significantly from those of normal males ( $p < .05$ ), whereas the comprehension scores of schizophrenic and normal females did not significantly differ.

The within portion of the analysis revealed a significant main effect for Type of Cartoon ( $F = 14.74$ ;  $df = 1/56$ ;  $p < .001$ ) and significant interactions for Sex  $\times$  Type of Cartoon ( $F = 5.13$ ;  $df = 1/56$ ;  $p < .05$ ) and Diagnosis  $\times$  Type of Cartoon ( $F = 10.13$ ;  $df = 1/56$ ;  $p < .005$ ). Across all subjects, greater comprehension was found on the normal than the abnormal cartoons. The Sex by Type of Cartoon interaction reflected a greater difference between normal and abnormal cartoons in males ( $p < .01$ ) than in females (difference not significant), as well as a greater difference between males and females on the normal cartoons ( $p < .01$ ) than on the abnormal cartoons (difference not significant). Most critical for the issues raised in this study was the significant Diagnosis by Type of Cartoon interaction, which may be seen in Figure 1. Further analysis of this interaction revealed that schizophrenics had significantly lower comprehension scores on the abnormal as compared to the normal cartoons ( $p < .01$ ), whereas the normal control subjects did not. The interaction also reflected the fact that while schizophrenics and controls did not



differ significantly on the normal cartoons, they did differ on the abnormal cartoons ( $p < .01$ ).

A more differentiated analysis of the effects of Type of Cartoon was performed in which five categories of cartoons (sex role irrelevant  $N = 5$ , normal

TABLE 4  
COMPREHENSION AND PREFERENCE SCORES FOR EACH CARTOON TYPE

COMPREHENSION AND PREFERENCE SCORES FOR DRAWINGS							
Group	N	Normal cartoons		Abnormal cartoons		All cartoons	
		$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Comprehension scores							
Schizophrenic							
Male	20	1.64	.47	1.27	.37	1.46	.42
Female	20	1.52	.32	1.35	.33	1.44	.33
Total	40	1.58	.40	1.31	.35	1.45	.38
Normal							
Male	10	1.96	.07	1.86	.18	1.91	.13
Female	10	1.51	.42	1.56	.26	1.54	.34
Total	20	1.73	.37	1.71	.27	1.72	.32
Preference scores							
Schizophrenic							
Male	20	1.51	.25	1.14	.38	1.33	.32
Female	20	1.36	.24	1.24	.43	1.32	.34
Total	40	1.44	.25	1.19	.40	1.32	.33
Normal							
Male	10	1.48	.32	1.28	.39	1.38	.36
Female	10	1.34	.28	1.31	.31	1.33	.30
Total	20	1.41	.30	1.29	.34	1.35	.32

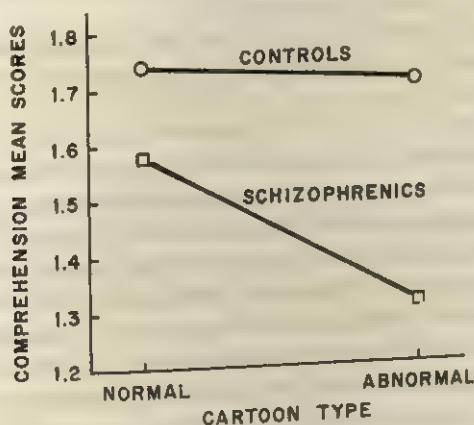


FIGURE 1  
COMPREHENSION SCORES: DIAGNOSIS  $\times$  CARTOON TYPE INTERACTION  
(Two Cartoon Types)

sex role  $N = 5$ , feminine men  $N = 5$ , masculine women  $N = 3$ , and ambiguous sex role  $N = 2$ ) rather than two were used. This analysis was conducted primarily to investigate the role reversal *versus* role alienation issue raised by McClelland and Watt (6). Support for the formulation that schizophrenics suffer from sex-role reversal could be obtained from the Humor Test in a three-way interaction indicating that male schizophrenics' comprehension scores were extremely low for those five cartoons depicting feminized men, whereas the female schizophrenics' comprehension scores were extremely low for those three cartoons depicting masculinized women. This  $2 \times 2 \times 5$  (Diagnosis  $\times$  Sex  $\times$  Type of Cartoon) analysis of variance revealed no significant differences between normal controls and schizophrenics for either the non-sex-role or the normal sex-role cartoons (see Figure 2). However, for each of the three types of abnormal cartoons—i.e., abnormal male, abnormal female, and ambiguous identity—the schizophrenics had significantly lower comprehension scores than the normal controls (all  $p$ s  $< .01$ ). No evidence was found ( $F < 1$ ) indicating that comprehension scores of male and female schizophrenics differed as a function of the category of cartoon.

The like-dislike scores of the groups are presented in Table 4. A repeated measures analysis of variance performed on these scores revealed only a significant main effect for Type of Cartoon ( $F = 16.05$ ;  $df = 1/56$ ;  $p < .001$ ) and a significant Sex by Type of Cartoon interaction ( $F = 5.31$ ;  $df = 1/56$ ;  $p < .05$ ). All subjects preferred the normal over the abnormal cartoons. The significant interaction reflected the fact that this preference was greater for males ( $p < .01$ ) than for females (not significant).

The mirth responses of the groups were also subjected to analyses, but no significant main effects or interactions were discovered.

#### D. DISCUSSION

Consistent with the conclusion of McClelland and Watt (6), the humor comprehension findings of the present study indicate that male and female schizophrenics are disturbed in their sex-role identification. Schizophrenics, as compared to the normal control group, displayed an inability to comprehend appropriately those cartoons depicting individuals engaged in abnormal or ambiguous sex roles. No such disability was found when schizophrenics were faced with the task of comprehending cartoons involving non-sex-role or normal sex-role behavior. That schizophrenics' inability to comprehend abnormal sex-role cartoons reflects a disturbance in sex-role identification is a conclusion stemming from earlier studies which have shown that the failure to comprehend cartoons with particular themes indicates that such themes re-

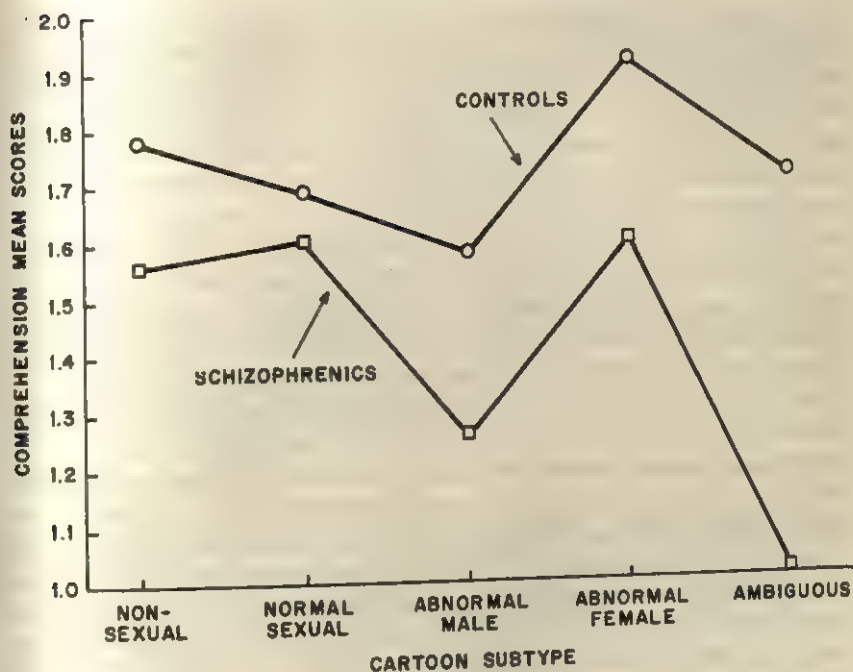


FIGURE 2  
COMPREHENSION SCORES: DIAGNOSIS  $\times$  CARTOON TYPE INTERACTION  
(FIVE CARTOON TYPES)

present a problem or conflict area. For example, Levine and Redlich (4) found that both psychiatric and highly intelligent normal individuals failed to comprehend cartoons which are ordinarily easily understood because the humor touched off strong emotional conflicts. The failure in dealing with these cartoons seemed to represent a defensive avoidance of the traumatic content.

Although the humor comprehension findings support the conclusion of McClelland and Watt (6), the schizophrenic-normal differences they reported on the Role Preference and Body Parts Satisfaction Tests were not replicated in the present study. The disparity in findings between this and the earlier study is most likely due to the fact that the schizophrenics in the McClelland and Watt study had been hospitalized for approximately 10 years, whereas those in the present study had been hospitalized for approximately four months.

This difference in length of hospitalization takes on importance in light of the McClelland and Watt finding that even schizophrenics hospitalized for 10

years do not differ from normals on an extremely overt measure of sex attitudes and interests. The view can be generated that it is the covertness or degree of subtlety of the sex-role alienation measure that will determine whether particular populations of schizophrenics will manifest sex-role alienation. Thus, as seen in the McClelland and Watt study, chronic and, therefore, more regressed schizophrenics display no such alienation on an extremely overt measure, but do on two more covert measures. It would appear that short-term hospitalized schizophrenics maintain that degree of intactness which permits them to call upon established patterns or ritualized acts, as can be found in expression of attitudes about parts of the body or in role playing, and thus not differ from normals on such measures. However, even the more psychologically intact schizophrenics represented by the short-term hospitalized population would display the sex-role alienation thought to characterize schizophrenics, providing a more subtle measure, such as the Humor Test, is employed.

In respect to this differential cognitive intactness with increasing periods of institutionalization argument, two specific hypotheses can be advanced: one dealing with the sampling phenomenon, and the second dealing with the effects of longer hospitalization *per se*. That different findings are obtained with schizophrenics institutionalized for four months as compared to those institutionalized for 10 years may be due to nothing more than a selective sampling procedure whereby those institutionalized for a long period represent a more severely disordered group. The alternative interpretation is that lengthening hospitalization results in a cognitive regression and/or little opportunity to practice sex roles, which results in a sex-role alienation even when gauged by relatively overt measures of sex-role identification. Selection between these two alternatives could be accomplished by studying severity of illness and then comparing short- and long-term hospitalized schizophrenics on measures of sex-role alienation that vary in degree of overtness.

The findings of the present study are in keeping with McClelland and Watt's contention that schizophrenics suffer from sex-role disturbance rather than sex-role reversal. The fact that there were no significant differences between male and female schizophrenics within comprehension of cartoons depicting inadequate males as opposed to cartoons depicting aggressive females suggests that the disturbance in sex identification is not simply one of sex reversal. As must always be the case when lack of findings is used to support a particular view, some caution is in order. It may be that the number of cartoons in the two categories was too small to permit those differences to emerge which would be indicative of this sex-role reversal which has been found in some studies. While the present study poses problems for further research, it does indicate that even



short-term hospitalized schizophrenics display disturbances in gender identification. It would appear that any comprehensive theory of schizophrenia must deal with the etiology of this disturbance and how this dynamic influences other facets of schizophrenic functioning.

#### REFERENCES

1. CHEEK, F. E. A serendipitous finding: Sex roles and schizophrenia. *J. Abn. & Soc. Psychol.*, 1964, **69**, 392-400.
2. HOLZBERG, J. Sex differences in schizophrenia. In H. Beigel (Ed.), *Advances in Sex Research*. New York: Harper & Row, 1963.
3. LEVINE, J., & ABELSON, R. Humor as a disturbing stimulus. *J. Gen. Psychol.*, 1959, **60**, 191-200.
4. LEVINE, J., & REDLICH, F. C. Some factors in the failure to understand humor. *Psychoanal. Quart.*, 1955, **24**, 560-572.
5. ———. Intellectual and emotional factors in the appreciation of humor. *J. Gen. Psychol.*, 1960, **62**, 25-35.
6. MCCLELLAND, D. C., & WATT, N. F. Sex-role alienation in schizophrenia. *J. Abn. Psychol.*, 1968, **73**, 226-239.
7. REDLICH, F. C., LEVINE, J., & SOHLER, T. P. A mirth response test: Preliminary report on a psychodiagnostic technique utilizing dynamics of humor. *Amer. J. Orthopsychiat.*, 1951, **21**, 717-731.
8. ZIGLER, E., LEVINE, J., & GOULD, L. Cognitive processes in the development of children's appreciation of humor. *Child Devel.*, 1966, **37**, 507-518.

*Psychology Service*  
*Veterans Administration Hospital*  
*West Haven, Connecticut 06516*



## RELATIONS AMONG SENSATION SEEKING AND SIMULATED AND BEHAVIORAL PERSONAL SPACE\*<sup>1</sup>

*Brigham Young University*

---

DARHL M. PEDERSEN<sup>2</sup>

---

### SUMMARY

Simulated and behavioral personal space measures and the Sensation Seeking Scale were administered to 20 males and 20 females. In the simulated measures Ss positioned male and female top profiles relative to a top profile representing themselves at each of nine equally spaced angles. For the behavioral measure Ss were approached by a male or female at each of the nine angles.

The female "swinger" had greater simulated personal space. Males who liked new and interesting experiences had greater side and diagonal simulated personal space, and females high on the same scale had a closer behavioral personal space toward males except at the sides. Males and females responded similarly to males and females approaching from various directions. However, there was no relation between front and side simulated personal space for males. Size of personal space relative to females tended to be highly correlated with size of personal space toward males. However, for male Ss there was little relation between front personal space toward males and females. For male Ss simulated personal space and behavioral personal space were highly correlated. For female Ss the two measures were almost completely unrelated.

---

### A. INTRODUCTION

Personal space refers to the area immediately surrounding a person which he regards as his own and which he does not normally like other people to penetrate (1, 2, 4, 8). The size of the space varies from person to person. And for a particular person it fluctuates depending upon location of the interaction, relationship to the other person, purpose of the interaction, characteristics of the individuals involved, as well as other factors.

---

\* Received in the Editorial Office on October 19, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This experiment was supported in part by a Faculty Research Fellowship granted by Brigham Young University.

<sup>2</sup> Requests for reprints should be sent to 1230A SFCL, Brigham Young University, Provo, Utah 84601.

Some attempts have been made to determine personality correlates of size of personal space. Patterson and Holmes (6) found that introverted females maintained a greater distance from a male interviewer than extroverted females. Ss high on manifest anxiety underestimated the distance between themselves and another person as compared with less anxious Ss (5). Schizophrenics had a larger personal space than nonschizophrenics (3).

Simulated tasks in which Ss manipulate figures or objects representing themselves and others have been related to actual personal space behavior. Dosey and Meisels (1) reported little relationship between a task of drawing silhouettes and actual behavior. Rawls, Trego, and McGaffey (7) found a correlation of only .34 between a simulated and a behavioral task.

The purpose of the present experiment was to (a) relate two scales on the Sensation Seeking Scale (9) to personal space, (b) relate a simulated personal space task to a behavioral task, and (c) investigate different directions of orientation toward other persons on simulated and behavioral personal space tasks.

## B. METHOD

### 1. Subjects

Ss were 20 males and 20 females enrolled in psychology classes at Brigham Young University.

### 2. Measuring Instruments

*a. Pedersen Personal Space Measure (PPSM).* The PPSM consisted of 18 pages on each of which was a stationary top profile of either a man or a woman, a four-inch line radiating from the profile at one of nine angles, and a gum-backed moveable top profile. The stationary profile was the same sex as S and was used to represent S. There were two sets of nine pages. For one set the moveable profile was male, and for the other it was female. In each set nine different angles were formed by the four-inch line—one on each of the nine pages. The angles were  $0^\circ$ ,  $22\frac{1}{2}^\circ$ ,  $45^\circ$ ,  $67\frac{1}{2}^\circ$ ,  $90^\circ$ ,  $112\frac{1}{2}^\circ$ ,  $135^\circ$ ,  $157\frac{1}{2}^\circ$ , and  $180^\circ$ , starting at S's right hand side and going to S's left hand side. At  $90^\circ$  the line extended directly in front of S.

Ss were instructed to remove the gum-backed profile representing another person and attach it on the line facing the S-profile so that the resulting distance was as close as possible so that S still felt comfortable; i.e., if the profile were to be moved closer, it would make S feel uncomfortable. The 18 pages were presented in random order.

*b. Pedersen Behavioral Personal Space Measure (PBPSM).* The PBPSM



was identical in form to the PPSM except that nine six-foot pieces of tape were placed on the floor radiating from a central point. Each piece of tape was marked at quarter-inch intervals. *S* stood facing the 90° tape. An unfamiliar person approached at each of the nine different angles until *S* indicated for each that he would feel uncomfortable if the person moved closer. Eighteen trials were presented in random order. For half of the trials the approaching person was a male, and for the other half the approaching person was a female. On each trial the score was the distance from the tip of *S*'s toe to the tip of the closest toe of the other person.

*c. Sensation Seeking Scale (SSS).* This instrument developed by Zuckerman (9) measured dimensions of sensation seeking. The Disinhibition subscale (Dis) consisted of items that expressed a hedonistic "playboy philosophy": heavy social drinking, variety in sexual partners, "wild parties," and gambling. The author suggests that an alternative label might be "Swinger." The Boredom Susceptibility subscale (BS) indicated in males dislike for repetitive experiences and routine work, preference for exciting people, and restlessness when things are unchanging. The BS scale was more clearly identified for males than females in the factor analytic construction of the test.

### 3. Procedure

The PPSM was given during a one-hour class period. The distance between each *S*-profile and the profile representing another person was recorded to the nearest millimeter. The PBPSM was administered individually during a 10-minute experimental session. The distance on each trial was recorded to the nearest quarter of an inch.

### 4. Data Analysis

The analysis of the data consisted of Pearson product-moment correlations to describe the following relationships: (*a*) the relation of sensation seeking to simulated and behavioral personal space requirements, (*b*) the relation among different directions of orientation to another person, (*c*) the relation of personal space relative to a male with personal space relative to a female for both simulated and behavioral tasks, and (*d*) the relation of simulated personal space to behavioral personal space for each of the directions utilized.

## C. RESULTS AND DISCUSSION

Four subsets of raw scores were obtained as follows: PPSM-male approaching, PPSM-female approaching, PBPSM-male approaching, and PBPSM-

female approaching. For each subset, scores were obtained for each of the nine angles of approach. A preliminary factor analysis demonstrated that the nine angles of orientation in each subset could be represented by three subscores: (a) Front Direction—distance at 90° angle, (b) Side Direction—mean of distances at 0° and 180°, and (c) Diagonal Direction—mean of distances at 22½°, 45°, 67½°, 112½°, 135°, and 157½°.

In addition, an average subscore for each subset was obtained by finding the mean distance for the nine angles. These four scores represented the average personal space relative to members of each sex for both the PPSM and the PBPSM. Finally, an overall average score was found for the PPSM and for the PBPSM. These scores were the means of the distances for all 18 angles in each personal space measure. Thus, instead of 36 personal space scores for each S, 18 scores were utilized. The description of the scores, the corresponding variable numbers, and the means and standard deviations are presented in Table 1.

### 1. *Relation of Sensation Seeking to Personal Space*

The correlation between the Sensation Seeking subscales of Disinhibition and Boredom Susceptibility was .59 for males (significant at the .005 level,

TABLE 1  
MEANS AND STANDARD DEVIATIONS OF PERSONAL SPACE VARIABLES  
AND SENSATION SEEKING SCALES

Variable number	Variable name	Mean		Standard Deviation	
		Males	Females	Males	Females
1	PPSM: M front	22.15	23.70	20.87	17.02
2	PPSM: M side	21.85	15.00	20.24	15.11
3	PPSM: M diagonal	22.54	19.96	13.91	12.16
4	PPSM: M average	22.34	19.27	14.64	12.50
5	PPSM: F front	10.35	18.50	7.84	8.95
6	PPSM: F side	14.58	11.30	18.25	8.50
7	PPSM: F diagonal	15.23	18.38	12.49	11.42
8	PPSM: F average	14.54	16.82	12.55	9.89
9	PPSM: average	18.44	18.05	12.93	11.02
10	PBPSM: M front	16.90	19.01	10.70	9.82
11	PBPSM: M side	17.20	15.75	9.42	8.89
12	PBPSM: M diagonal	17.17	16.37	9.03	8.26
13	PBPSM: M average	17.15	16.53	8.88	8.22
14	PBPSM: F front	13.43	17.19	10.64	9.68
15	PBPSM: F side	16.60	14.83	9.66	7.81
16	PBPSM: F diagonal	14.66	16.11	10.78	7.63
17	PBPSM: F average	14.95	15.95	10.22	7.69
18	PBPSM: average	16.05	16.24	9.09	7.72
19	SSS: Dis	2.90	2.40	3.04	2.21
20	SSS: BS	6.50	7.10	3.07	2.53

Note: PPSM = Pedersen Personal Space Measure; PBPSM = Pedersen Behavioral Personal Space Measure; SSS = Sensation Seeking Scale; Dis = Disinhibition subscale; BS = Boredom Susceptibility subscale. PPSM units are millimeters, and PBPSM units are inches.

one-tailed) and .15 for females (nonsignificant). This finding may be explained by the statement by Zuckerman (9) that the BS scale was ill defined for females. The subscales are independent for females, but not for males.

Table 2 presents the correlations of the two subscales with the personal

TABLE 2  
CORRELATIONS OF SENSATION SEEKING SCALES WITH  
PERSONAL SPACE MEASURES

Variable number	Sensation Seeking Scale			
	19. Disinhibition	20. Boredom Susceptibility		
	Males	Females	Males	Females
1	.16	.40*	.02	— .34
2	.26	.25	.41*	— .32
3	.14	.53**	.36	— .34
4	.20	.47*	.36	— .36
5	— .04	.20	.13	— .21
6	.22	.35	.47*	— .09
7	.13	.56**	.46*	— .16
8	.15	.52**	.47*	— .16
9	.18	.50*	.43*	— .28
10	— .12	— .24	.01	— .53**
11	— .08	— .28	.19	— .17
12	— .07	— .29	.18	— .44*
13	— .08	— .29	.17	— .40*
14	— .09	— .35	.20	— .30
15	.06	— .20	.33	— .15
16	— .03	— .24	.32	— .29
17	— .02	— .25	.32	— .26
18	— .05	— .28	.26	— .35

\*  $p < .05$  (one-tailed).

\*\*  $p < .01$  (one-tailed).

space measures. The patterns of intercorrelations were distinct for males and females. For female Ss, Disinhibition correlated with personal space on the simulated task. It did not relate to females' behavioral personal space or to simulated or behavioral personal space for males. The female "swinger" tended to place the top profile of other persons farther away from her. This was particularly true of male profiles placed in front and diagonal orientations, and of female profiles placed in diagonal orientations. The correlation with the overall average on the PPSM was .50. In relation to real people the female swinger did not have a larger personal space. In fact the trend, although the correlations were not significant, was for her to get closer. The male with a "playboy philosophy" did not respond to others conceptually or behaviorally with a closer or more distant interpersonal proximity.

Whereas Disinhibition correlated with PPSM for females and not males, Boredom Susceptibility correlated with the PPSM for males and not females. The males who liked new and interesting activities and people preferred more

distance at the side orientations toward males, and the side and diagonal directions toward females. They did not indicate a greater distance frontally than people low on BS. All of the PPSM correlations with the BS scale were significant or close to significance for males except those involving the front direction. It is unknown why a closer front orientation was not also indicated. An explanation will have to await additional research regarding the personality dynamics of males high on the BS scale. The BS scale did not correlate significantly with any of the behavioral personal space measures for males.

Females with high susceptibility to boredom tended to prefer males closer in the behavioral task both at the front and diagonal directions. In fact negative correlations prevailed for both the PPSM and PBPSM. Apparently, females who preferred exciting people preferred males closer except at the side.

## 2. *Relation Among Different Directions of Orientation to Another Person*

In the preceding discussion some selectivity was obtained in personal spacing depending upon the direction of orientation. Table 3 presents results that assess the similarity of the three directions of orientation.

TABLE 3  
CORRELATIONS AMONG THREE DIRECTIONS OF APPROACH BY APPROACHING  
MALES AND FEMALES ON THE PPSM AND THE PBPSM  
FOR MALE AND FEMALE Ss

Measure	Front Males	with side Females	Front Males	with diagonal Females	Side with Males	diagonal Females
PPSM:						
Approaching males	.35	.71**	.55*	.89**	.86**	.72**
PBPSM:						
Approaching females	.20	.56*	.54*	.74**	.91**	.76**
PBPSM:						
Approaching males	.64**	.63**	.79**	.85**	.87**	.89**
PBPSM:						
Approaching females	.77**	.83**	.91**	.91**	.89**	.91**

Note: PPSM = Pedersen Personal Space Measure; PBPSM = Pedersen Behavioral Personal Space Measure.

\*  $p < .01$  (one-tailed).

\*\*  $p < .001$  (one-tailed).

A comparison of the magnitudes of the correlations among the three directions of orientation shows that the side and diagonal orientations of Ss tend to be most similar. Front and diagonal personal space distances are almost as highly correlated. And front and side orientations are most dissimilar in personal space requirements.



There was a tendency for Ss to exhibit greater consistency among the three orientations on the behavioral task than on the simulated task. The relative magnitudes of correlations were the same for females as for males when approached by males and by females for the three directions. For example, the correlation of front orientation with side orientation on the PBPSM was larger for males approached by females than for males approached by males. It was also larger for females approached by females than for females approached by males. An examination of all six such comparisons does not reveal a single inversion. Females tended to respond to the various angles of approach the same way that males did to both approaching males and females.

For the PBPSM the correlations among the three directions are larger for approaching females than for approaching males. Thus, on the behavioral measure, greater consistency was found in personal space specifications for both males and females when approached by females than when approached by males. No such generalization is possible for the simulated task. The consistency of personal spacing appeared to be about the same for males and females approached by males and females, except for the relation between front and side personal space requirements. Both males and females tended to be more similar in the front and side personal space requirements when approached by males than when approached by females.

For males, there was no significant correlation between front and side personal space responses when approached by males or females. On the simulated task males responded quite distinctly in relation to either males or females at the front as compared to the side. This fact is corroborated in Table 2, where for males BS correlated quite highly with side and diagonal measures relative to both approaching males and females, but not with the front measures. In general, however, both males and females tended to be fairly consistent in their personal space responses toward males and females on the simulated and behavioral tasks for the three directions of orientation. Ss with larger personal space requirements tended to be uniformly larger for the three directions of approach. Conversely, Ss with smaller personal spaces tended to be smaller for all three directions.

### *3. Relation of Personal Space Relative to a Male with Personal Space Relative to a Female*

Table 4 gives correlations between corresponding male and female approaching items on the PPSM and the PBPSM. The correlations indicated that the Ss tended to respond similarly to members of the opposite sex and members of

TABLE 4  
CORRELATIONS BETWEEN CORRESPONDING MALE AND FEMALE APPROACHING SCORES  
ON THE PPSM AND PBPSM FOR MALE AND FEMALE Ss

Measure	Variables	Males	Females
PPSM	1 with 5	.02	.80**
	2 with 6	.91**	.81**
	3 with 7	.85**	.95**
	4 with 8	.81**	.94**
PBPSM	10 with 14	.59*	.73**
	11 with 15	.79**	.92**
	12 with 16	.83**	.85**
	13 with 17	.81**	.88**

Note: PPSM = Pedersen Personal Space Measure; PBPSM = Pedersen Behavioral Personal Space Measure.

\*  $p < .005$  (one tailed).

\*\*  $p < .001$  (one-tailed).

the same sex with respect to personal space on both the simulated and behavioral tasks. The degree of relationship appeared to be about the same for the various directions and for the two measures except for the relation between front approaching males and front approaching females on both the PPSM and the PBPSM for male Ss. This finding is consistent with earlier results which indicated that the response of male Ss to the front direction was distinct from their other personal space measures. There was no relationship between male Ss' personal space toward males and toward females in the face-to-face orientation on the simulated task. And on the behavioral task, although the correlation was significant, it was noticeably less than the other correlations between corresponding tasks relative to males and females.

#### 4. Relation of Simulated Personal Space to Behavioral Personal Space

The PPSM and the PBPSM were devised so that the tasks would be as similar as possible on the two measures. There were nine pairs of corresponding scores obtained from the two measures. The correlations between these pairs of corresponding scores for male and female Ss are presented in Table 5. These correlations indicated the extent to which behavioral personal space may be predicted from simulated personal space for the tasks utilized.

An inspection of the correlations discloses that the largest degree of correspondence in personal space between simulated and behavioral measures was between pairs of female approaching items for male Ss. Males tended to respond toward females the same behaviorally as they did on the simulated task. Although, the correlations are not as large, there was a tendency for the males also to be fairly consistent in their responses toward males on the two measures. The notable exception is the correlation between frontally approach-

TABLE 5  
CORRELATIONS BETWEEN CORRESPONDING TASKS ON THE PPSM  
AND THE PBPSM FOR MALE AND FEMALE Ss

Variables	Males	Females
1 with 10	.11	.32
2 with 11	.51**	.41*
3 with 12	.50*	.33
4 with 13	.45*	.35
5 with 14	.69***	.10
6 with 15	.69***	.83***
7 with 16	.80***	.16
8 with 17	.77***	.24
9 with 18	.62**	.30

Note: PPSM = Pedersen Personal Space Measure; PBPSM = Pedersen Behavioral Personal Space Measure.

\*  $p < .05$  (one-tailed).

\*\*  $p < .01$  (one-tailed).

\*\*\*  $p < .001$  (one-tailed).

ing males on the two measures. This correlation is not significant. Again male Ss responded quite distinctly with respect to frontally oriented people. Here the uniqueness of the response involved the personal space toward frontally oriented males. The relationship between front orientation responses toward females on the two measures is consistent with the relationships between corresponding tasks. The correlation between the overall average on the PPSM and the overall average on the PBPSM is .62.

Contrary to the results for males, the responses of females on the two measures were almost totally unrelated. The exceptions were the correlations between males approaching from the side and between females approaching from the side for the two measures. In other words, female Ss tended to establish the same personal space toward males and toward females at the sides on the PPSM and the PBPSM. Otherwise personal space indicated on the simulated measure was unrelated to personal space on the behavioral measure for females. The correlation between overall average personal space scores for the two measures was a nonsignificant .30.

It is evident that personal space must be considered separately for males and females. They responded differentially in personal space behavior even when they had similar personality characteristics. Males responded distinctly in setting personal space toward other people in a direct face-to-face orientation on a simulated task. Not only were their front direction responses unrelated to side and diagonal directions toward both males and females, but also their front direction responses toward males and females were unrelated. Furthermore, their simulated front personal space toward males was unre-

lated to their behavioral front personal space toward males, even though on all other measures males' simulated personal space was highly related to their behavioral personal space. On the other hand, for females, simulated personal space in all directions of orientation except the side direction were unrelated to behavioral personal space.

Otherwise, personal space responses were similar for males and females. They tended to establish the same personal space toward others, irrespective of the direction of orientation, both in simulated and behavioral tasks. They tended to establish the same personal space toward males and females on comparable tasks.

### REFERENCES

1. DOSEY, M. A., & MEISELS, M. Personal space and self protection. *J. Personal. & Soc. Psychol.*, 1969, **11**, 93-97.
2. HOROWITZ, M. J. Body image. *Arch. Gen. Psychiat.*, 1966, **14**, 456-460.
3. HOROWITZ, M. J., DUFF, D. F., & STRATTON, L. O. Body-buffer zone: Exploration of personal space. *Arch. Gen. Psychiat.*, 1964, **11**, 651-656.
4. LITTLE, K. B. Personal space. *J. Exper. Soc. Psychol.*, 1965, **1**, 237-247.
5. LUFT, J. On nonverbal interaction. *J. of Psychol.*, 1966, **63**, 261-268.
6. PATTERSON, M., & HOLMES, D. S. Social interaction correlates of the MPI extroversion-introversion scale. *Amer. Psychol.*, 1966, **21**, 724-725.
7. RAWLS, J. R., TREGO, R. E., & MCGAFFEY, C. N. A comparison of personal space measures. Report of NASA Grant NGR-44-009-008, Institute of Behavioral Research, Texas Christian University, Fort Worth, October, 1968.
8. SOMMER, R. Studies in personal space. *Sociometry*, 1959, **22**, 247-260.
9. ZUCKERMAN, M. A new form (IV) of the Sensation Seeking Scale. Unpublished manuscript, University of Delaware, Newark, 1969.

*Department of Psychology*  
*Brigham Young University*  
*Provo, Utah 84601*



## SOME CORRELATES OF THE LEVEL OF CONSTRAINT IN A SYSTEM OF SOCIAL ATTITUDES\*<sup>1</sup>

*Department of Psychology, Oklahoma State University*

---

WILLIAM W. RAMBO, WARREN H. JONES, AND PHILLIP D. FINNEY<sup>2</sup>

---

### SUMMARY

Variables thought to be associated with the magnitude of interrelationship observed among the elements in a system of social attitudes were investigated. Subjects ( $N = 753$ ) were asked to respond to an inventory requesting self-report and background data. Several weeks later, Ss were given two forms of the Social Attitude Scale, a scale tapping the liberalism-conservatism domain. The results indicated that those Ss who reported recent changes in religious, moral, or political attitudes displayed a more highly interconnected system of liberal-conservative attitudes than those who did not. Variables relating to active participation in social affairs tended to be associated with higher levels of attitude system organization. Sex differences were observed, with men displaying higher levels of system organization than women. This relationship held even when participation rates were taken into consideration.

### A. INTRODUCTION

The attitude that an individual holds regarding a single social event is quite likely related to the attitudes he holds concerning other events. As a result, rather than functioning as independent elements, many social attitudes are better seen as members of a system that possesses an organizational structure. Researchers have begun to turn their attention toward these structural properties, and several studies have been reported (1, 3, 4) the results of which provide interesting and provocative indications of the behavioral significance of this area of inquiry.

Converse (1), for example, has investigated the level of constraint found in a system of political attitudes. The term "constraint" is used to refer to the extent to which the individual elements in a system show evidence of interrela-

---

\* Received in the Editorial Office on October 20, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This study has been supported by the Oklahoma State University Research Foundation, Marvin T. Edmison, Director.

<sup>2</sup> The authors wish to acknowledge the generous assistance received from Mr. William V. Accola, of the University Computing Center, and Mr. Larry S. Tomlinson, of the Research Foundation.

tionship, and it is most directly measured in terms of the magnitude of the intercorrelations observed among a set of items that have been drawn from a particular domain. With increasing levels of constraint, one may begin to think of the system as an integrated ideology; at decreasing levels, the items begin to appear as a fragmented collection of attitude elements. Converse's data demonstrate a broad distribution of individual differences in the degree of interrelationship among belief elements in the system studied, with a relatively small group of Ss whose attitudes were internally organized to the extent that might justify thinking of them as a political ideology. In addition, his data suggest that a linear function may describe the relationship between the level of constraint in the system and the amount of information the individual possesses regarding social and political events. Hence, as the individual acquires increasingly higher levels of information, there appears to be a corresponding change in the internal structure of the attitude system.

Results of this type tend to broaden our view of the behavioral implications of the term "attitude." Attitude change, for example, might well be described as a revision of the structure of the system, as well as in the more conventional terms of a linear movement of a point along a scale of affectivity. The revision of an individual's attitudes in this instance is seen as a restructuring of the system in which the range of the domain may broaden or constrict, its dimensional characteristics may become more or less complex, the level of constraint may strengthen or weaken, and so forth.

The present study represents an attempt to consider a set of variables that may be related to the level of constraint observed in a system of social attitudes found in the liberalism-conservatism domain. The study is an exploratory undertaking with its principal purpose being to point to possible lines of research that may improve our understanding of this structural attribute of social attitude systems.

#### B. PROCEDURE

The Ss in the present study were 753 undergraduates enrolled in a course in introductory psychology at Oklahoma State University. Eighty-six percent of the group was between the ages of 17 and 19, 56 percent was female, 43 percent indicated a preference for the Democratic Party, and the sample included Ss from all the major subject matter divisions on the campus.

The study was divided into two phases: the first conducted during a regular class period. At this time Ss were given a 50 item multiple choice test containing questions, the answers to which require information concerning current affairs. Items tapped Ss knowledge concerning political figures, organizations,

legislative and judicial decisions, political processes and traditions, and political philosophy. In addition, Ss were asked to fill in a form requesting background information. In was the information obtained from this form that served as the independent variables in the study. The S was asked for information concerning age, sex, subject matter concentration, and level of participation in political affairs. In addition, the S was asked to indicate on a seven category scale the extent to which his attitudes had changed within the last year. Ss reported perceived attitude change in areas of politics, religion, and morality.

The second phase of the study was carried out from two to three weeks after the administration of the first set of materials. Effort was made to dissociate one phase of the study from the other; hence, no reference was made to the first session, and personnel involved in this phase were not those involved in the first. Ss were asked to volunteer for an out-of-class experiment and were scheduled to report to the experimental room at a convenient time.

When Ss reported for the experiment they were given Forms A and L of the Social Attitude Scale (2). Each form of the scale consists of 44 Likert-type items which purportedly tap attitudinal elements found in the liberalism-conservatism domain. Items deal with such considerations as the nature of man, the nature of social order, and concepts dealing with social tradition and social change. This scale has been constructed on the basis of an assumption that attitudes toward more specific social referents—i.e. politics, religion, etc.—in some fashion derive from these more basic attitudes tapped by this instrument.

### C. RESULTS

The general procedure followed in the analysis was to create independent variable categories on the basis of S's response to an item on the background information schedule. Then, for each of these categories, a matrix of interitem contingency coefficients was computed for the items of the liberalism-conservatism scale. The magnitude of this coefficient represented the degree of constraint existing between a pair of items, and the distribution of magnitudes in a matrix was employed to represent the level of constraint in the entire system. Coefficients were arranged in order of magnitude for each experimental group, and a median test (5) was computed in order to determine whether the distributions differed significantly.

The first variable considered for analysis dealt with the S's report on the extent to which his attitudes had shifted during the recent past. Those reporting a substantial change were placed in one group; those reporting little or no change were placed in a second. For each group, a matrix of contingency coefficients was computed for the items of the Social Attitude Scale. Table 1 presents

TABLE 1  
AVERAGE INTERITEM CONTINGENCY COEFFICIENTS OBTAINED FROM Ss  
REPORTING SMALL OR LARGE CHANGES IN ATTITUDES

Social Attitude Scale form	Area of reported change		
	Political	Religion	Moral
Small change			
Form A	.274	.275	.278
Form L	.293	.299	.299
Large change			
Form A	.395	.340	.342
Form L	.396	.347	.358

mean coefficients for each group, and an examination of these values reveals that in all instances those Ss who reported a recent change in attitudes presented average interitem correlations that were larger than those obtained from Ss who reported little or no change in attitude. These differences held consistently for both forms of the scale. In addition, median tests carried out between the distribution of coefficients generated by each pair of experimental groups were all significant beyond the .01 level. Thus, it would appear that the internal structure of the broad system of social attitudes known as the liberalism-conservatism domain was more highly constrained for those Ss who reported a recent shift in political, moral, or religious attitudes.

At this point a question may be raised concerning the extent to which those Ss who reported a recent change in, say, political attitudes, may be the same Ss who reported revisions of religious or moral attitudes. However, contingency coefficients computed for each pairing of the three attitude domains considered were consistently low, ranging from .16 to .21; hence, although there was some tendency for those who indicated a recent change in one set of attitudes to report changes in the second, this relationship was quite low. Finally, reported attitude change did not appear to be related to the amount of information held by the S. For each of the three areas—i.e., political, religious, and moral—there was no significant difference between average information scores obtained by those whose attitudes had or had not recently undergone revision.

Matrices of contingency coefficients were next generated from the groups that were formed on the basis of each of the remaining background information items. Median tests revealed that the distribution of item intercorrelations was significantly higher for those belonging to a political party or who had registered to vote as compared to those who had not, for those who had actively participated in a political campaign as compared to those who had not, and for men as compared with women. This difference between the sexes remained



even when distributions were computed separately for men and women who had been active in political campaigns. These groups differed significantly from comparable groups of nonparticipants, but men participants still displayed higher constraint levels than did women participants ( $p < .01$ ).

When the data were analyzed with regard to the area of subject matter concentration, there were no significant differences observed between those who majored in the biological-physical sciences, social sciences, or education, or those who had not declared a major area of study. Those who were enrolled in the business college did present a significantly lower set of coefficients than those in the other groups. Finally, those registered Republican did not significantly differ from those registered Democrat. All differences obtained were consistent across both forms of the attitude scale.

#### D. DISCUSSION

The fact that a system of social attitudes appears rather stable, not having changed over the recent past, does not necessarily imply a well formed structure that is internally anchored by a network of interrelationships. On the contrary, the results of the present study indicate that it was in the group of Ss who reported recent changes in attitude that the liberalism-conservatism system was most highly constrained. Whether this higher degree of internal integration is a consequence of or a prerequisite for attitude change is a question that must await further research. To think of attitude change in terms of a restructuring of the interrelationships found among a set of elements seems quite reasonable; however, the notion that attitude change may, under certain circumstances, be facilitated by a pre-existing structure is a concept that should also be explored further.

Although any hypotheses drawn from this study must be highly tentative, the results would appear to indicate that higher levels of internal constraint tend to be found in association with those variables that imply active involvement, as opposed to variables that refer to affiliational status. For example, registering to vote or participation in a political campaign were found to be related to higher levels of attitude integration, whereas those Ss registered Republican did not differ significantly from those registered Democrat. In addition, subject matter concentration, with the exception of business majors, did not differentiate groups on the basis of this structural attribute of attitude systems. However, sex did influence the results, with men tending to present a distribution of interitem correlations that was significantly higher than was found in association with the women Ss. That these sex differences remained even when participation levels were taken into consideration might

point to the operation of a third variable which mediates this difference, or it might reflect one of the consequences of culturally imposed sex roles which may in one fashion or another discourage the development of highly constrained systems of social attitudes in those women who are represented by the present sample. In either case, the present research has established that the internal organization of the liberalism-conservatism system does vary as a function of several of the variables considered.

#### REFERENCES

1. CONVERSE, P. The nature of belief systems in mass publics. In D. Apter (Ed.), *Ideology and Discontent*. New York: Free Press, 1964.
2. RAMBO, W. W. The measurement of broad spectrum social attitudes: Liberalism-conservatism. *Percept. & Motor Skills*, in press.
3. ROKEACH, M. *The Open and Closed Mind*. New York: Basic Books, 1960.
4. ———. *Beliefs, Attitudes, and Values*. San Francisco: Jossey-Bass, 1968.
5. SIEGEL, S. *Nonparametric Statistics*. New York: McGraw-Hill, 1956.

*Department of Psychology*  
*Oklahoma State University*  
*Stillwater, Oklahoma 74074*

## REACTIONS TO AGGRESSION-RELATED STIMULI FOLLOWING REINFORCEMENT OF AGGRESSION\*<sup>1</sup>

*University of Missouri*

---

RUSSELL G. GEEN AND DAVID STONNER

---

### SUMMARY

Twenty-four male subjects were reinforced for aggressing with verbal approval from the experimenter, or not reinforced, then presented nine verbs previously scaled for aggressiveness associations as cues for further aggressive responses. Subjects who received approval from the experimenter increased their level of aggressiveness over reinforced trials, whereas nonreinforced subjects did not. Reinforced subjects also reacted with greater aggressiveness to words having strongly aggressive connotations than did nonreinforced subjects, but were no more aggressive than the latter when words having weaker connotations for aggression served as eliciting cues.

---

### A. INTRODUCTION

Geen and Stonner (8) reported an experiment in which male subjects were either reinforced or not reinforced by experimenter approval for delivering shocks to an experimental confederate in response to a signal light. Subjects who received reinforcement were found subsequently to deliver significantly more intense shocks to the same confederate than did nonreinforced subjects when the signal for giving the shock was the presentation of a word having strongly aggressive connotations. When the signal was a word having weak aggressive connotations, previously reinforced subjects gave no stronger shocks than subjects who had received no reinforcers. The data were interpreted as showing that individual differences in aggressiveness habit strength acquired through differential reinforcement produce differences in aggressive behavior only when stimuli adequate to elicit aggressive responses are present in the situation. The findings were therefore considered relevant to Berkowitz's (3) hypothesis regarding the importance of aggressive cues in aggressive behavior.

---

\* Received in the Editorial Office on November 6, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This study was supported by Grant GS-2748 from the National Science Foundation to the first author.

The present study was designed as an attempt to test whether the Geen-Stonner findings could be extended to a broader range of verbal stimuli than the ones used in the original experiment. Geen and Stonner (8) used 11 verbs as stimuli, taken from a list originally formulated by Gellermann (9); of these 11 verbs, six were taken to be generally nonaggressive in connotative meaning (*wash, travel, relax, walk, sleep, listen*), whereas five were considered to be relatively aggressive (*choke, massacre, murder, stab, torture*). In the present study, however, we wished to attain greater precision in the degree to which any verb satisfied a criterion for connotative "aggressive" meaning. In order to do this, we had first to select a large number of verbs falling into roughly defined categories labelled "aggressive" and "nonaggressive," and then to apply psychological scaling to some sample of words drawn from this pool. This was done, and nine verbs falling along a continuous scale representing aggressiveness of connotative meaning were used in the subsequent experiment.

The experiment reported here involved two groups of subjects, one of which was reinforced with verbal approval by the experimenter for shocking another person, and one of which was not reinforced for giving shocks. After the reinforcement or nonreinforcement for aggression had occurred, each subject was again asked to shock the other person whenever a word was given as a signal for the response. The words presented to the subject were the nine scaled verbs described above. We expected that subjects who had been reinforced for aggressing would be giving shocks of greater intensity at the end of the reinforcement period than would subjects not receiving reinforcement. We further expected that reinforced subjects would give shocks of greater intensity when words having strong connotations for aggression served as signals than when words of less strong connotations for aggression were used. Finally we expected that the gradient for aggression across the dimension of aggressiveness connotation would be higher and steeper for reinforced subjects than for nonreinforced subjects. In forming this latter prediction we were guided by Miller's (12) notion of "liberalized" S-R theory, and sought accordingly to describe the predicted outcomes as a case of stimulus generalization for aggressive responses. To do this we assumed that those of our nine verbs that ranked at the high end of the scale of aggressiveness connotations were more related to aggression (such as that which transpired in the first part of the experiment) than words lower on the same scale. Thus the words at the high end of the scale should elicit more aggression when they are used subsequently as cues for aggressive reactions than would words lower on the scale. In other words we predicted a generalization gradient for aggressive responses across the dimen-



sion of aggressiveness connotations of the eliciting words. Such liberalized extensions of the construct of generalization to aggression have been utilized in several previously reported studies (e.g., 2, 7). Furthermore, we also expected that the generalization gradient for reinforced subjects would be both higher and steeper than that for nonreinforced subjects. It should be higher because the habit strength associated with the aggressive response to the stimuli in the first part of the experiment should be greater in reinforced than in nonreinforced subjects; hence the response potential should be similarly greater for these subjects when generalized stimuli are used. The steeper gradient for reinforced subjects is predicted from the finding of the Geen and Stonner (8) study that when nonaggressive words are used as stimuli, reinforced subjects are no more aggressive than nonreinforced. Thus the two gradients should converge at lower values on the aggressiveness connotation scale.

## B. METHOD

### 1. *Subjects*

Twenty-seven males drawn from introductory psychology classes served as subjects and received points toward their final grades for participating. Of these subjects, three in the Reinforced condition were able after the experiment to verbalize the reinforcement contingency and were dropped from the sample to be replaced by other males drawn from the same population. None of the 12 subjects comprising the final reinforced group expressed awareness of the reinforcement contingencies.

### 2. *Procedure*

Complete details of the reinforcement procedure have previously been reported elsewhere (7, 8). Each subject was first instructed to deliver an electric shock to an experimental confederate, who posed as a subject, each time the latter made an error in a learning task. The subject was free to choose from among 10 intensities of shock. After 10 trials on which the subject gave shocks without comment from the experimenter, a series followed during which the experimenter gave half the subjects verbal approval each time the latter gave shocks of greater intensity than the average of the first 10 (Reinforced condition). The other half of the subjects were given no approval for shocking (Nonreinforced condition). Subjects in the Reinforced and Nonreinforced conditions were matched in terms of the intensity of shocks given during the baseline period: each nonreinforced subject was matched to a reinforced subject who had previously given shocks of approximately the same magnitude.

For this purpose, subjects were treated as equal if their baseline intensities fell within a corresponding tenth of the range of mean intensities; e.g., 1.0-1.9, 2.0-2.9, etc. Each nonreinforced subject was allowed to give as many shocks to the confederate as his reinforced counterpart had given.

Following the reinforcement period, each subject was told that the second part of the experiment was concerned with effects of shock on the encoding of stimuli. The subject was therefore instructed to shock the confederate each time one of nine words appeared on a screen, supposedly to interfere with the confederate's encoding of the verbal stimulus. The nine words were the verbs that had previously been scaled (see below) along an "aggressiveness" dimension. The subject was given freedom to choose the intensity of the shock given on each of the nine trials. The order of presentation of the nine verbs was randomized across subjects within reinforcement conditions. At the conclusion of the second shock series, the experimenter explained that the experiment was over. The subject was interviewed, so that any failures in the experimental deception could be detected, and then dismissed.

### 3. *Selection of Stimulus Words*

Prior to the experiment, 100 male judges who had no other connection with the study were asked to generate lists of verbs consisting of (a) the 20 most violent they could think of, (b) the 20 least violent, and (c) 20 that had moderately violent associations. The eight verbs reported most frequently on each list were chosen for further use. From these, three words listed as highly violent were found to correspond in frequency of usage to three words from each of the other lists according to the Thorndike-Lorge lists. The nine words that emerged from this pretesting were then scaled by 50 naive male judges for connotations of violence according to the method of pair comparisons (11). Internal consistency assumptions of Case V were met. The resulting scale values of the nine verbs, with a low score indicating highly violent connotation, were as follows: *annihilate* (.00), *kill* (.21), *destroy* (.27), *mangle* (.80), *beat* (1.09), *tear* (1.90), *reduce* (2.25), *walk* (2.56), and *evolve* (2.84).

## C. RESULTS AND DISCUSSION

### 1. *Effects of Reinforcement*

Some reinforced subjects gave a few shocks that were below their baseline level and thus were not reinforced on these trials. The total number of shocks given by reinforced subjects (and their matched controls) thus differed from subject to subject. The data on shocks given by each subject were therefore

TABLE 1

MEAN INTENSITIES OF SHOCKS GIVEN BY REINFORCED AND NONREINFORCED SUBJECTS

Condition	Tenths of total trials									
	1	2	3	4	5	6	7	8	9	10
Reinforced	4.00	5.20	6.75	6.77	6.75	7.33	6.10	6.75	7.70	7.90
Nonreinforced	4.20	4.80	4.50	5.10	5.60	4.15	4.25	4.50	5.00	3.00

Vincentized, and a mean intensity of each tenth of all shocks given was calculated. These mean intensities for reinforced and nonreinforced subjects are shown in Table 1. An analysis of variance shows a significant main effect for Reinforcement ( $F = 6.68$ ,  $df = 1/22$ ,  $p < .025$ ) and a Reinforcement  $\times$  Trials interaction ( $F = 2.58$ ,  $df = 9/98$ ,  $p < .025$ ). The interaction effect indicates that subjects who were reinforced for shocking became progressively more aggressive, whereas matched controls who were not reinforced did not increase in their level of aggression over trials.

## 2. Reactions to Verbal Stimuli

Figure 1 presents the mean intensities of shocks given by reinforced and nonreinforced subjects when the nine verbs were eliciting stimuli. The abscissa represents an interval scale along which verbs may be located according to their aggressive association value. The scale value of each of the nine words used in the present study is indicated. Reinforced subjects gave shocks of greater intensity in response to words having strong aggressive connotations than to words of intermediate aggressive connotation, and still less intense shocks to words at the lower end of the scale. Nonreinforced subjects showed less of a tendency to decrease shock intensities as the eliciting words became less violent in association. An analysis of variance of these data reveals a significant main effect for Words ( $F = 10.61$ ,  $df = 8/176$ ,  $p < .001$ ) and a significant Reinforcement by Words interaction ( $F = 3.68$ ,  $df = 8/176$ ,  $p < .001$ ).

An additional test was made to determine the significance of the difference in slopes for the Reinforced and Nonreinforced groups. The Reinforcement  $\times$  Words interaction was broken down into a linear component and pooled non-linear component by means of an Alexander trend analysis (10). The resulting Between Groups linear effect was found to account for most of the interaction variance ( $F = 28.57$ ,  $df = 1/176$ ,  $p < .001$ ) indicating that the difference in the gradients of the two groups is a linear one. Finally, the deviation of each of the two gradients from zero slope was tested. The results of this test revealed that the linear regression coefficient for the gradient generated by reinforced

subjects differed significantly from zero ( $t = 6.46, df = 7, p < .01$ ), while that for the gradient generated by nonreinforced subjects fell just short of significant difference from zero ( $t = 2.30, df = 7, p = .07$ ). We may conclude, therefore, that our expectation concerning a steeper gradient of generalization across aggressive words for reinforced subjects than for nonreinforced subjects was supported.

We should recall at this point that none of the subjects who were reinforced for aggressing verbalized the reinforcement contingencies; all can therefore be assumed to have been unaware of being reinforced. Thus we may conclude that the results of the present study are not due to the reinforced subjects' reacting to demand characteristics of the situation created by the experimenter's approval of aggression. Furthermore, explanation of the actions of reinforced subjects on the basis of manding behavior fails to account for the results of the generalization phase of the experiment. If reaction to experi-

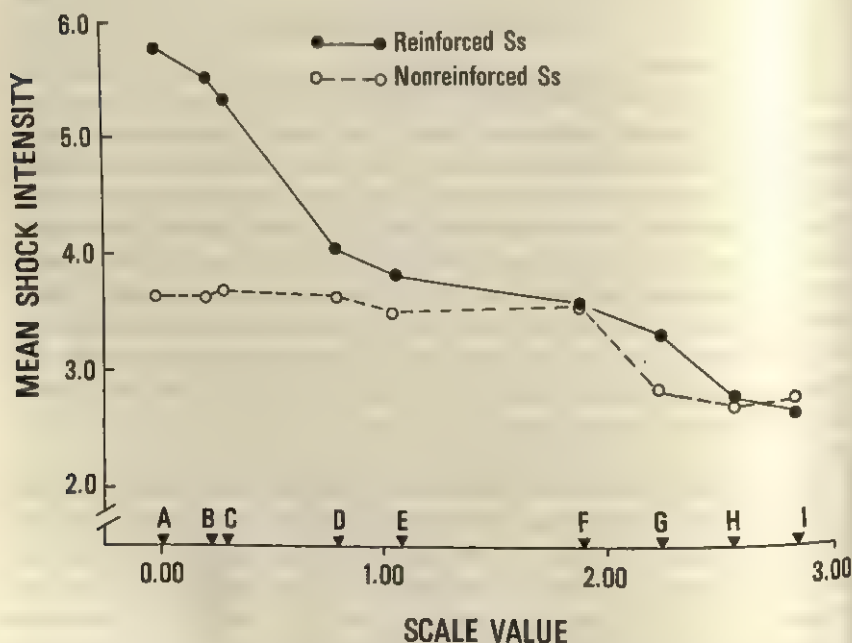


FIGURE 1

MEAN SHOCK INTENSITIES USED BY SUBJECTS IN RESPONSE TO CUE WORDS

Points on abscissa noted by alphabetic characters represent scale values of the nine verbal stimuli: A = annihilate, B = kill, C = destroy, D = mangle, E = beat, F = tear, G = reduce, H = walk, I = evolve.



menter approval alone accounts for the aggressiveness of reinforced subjects during training, these subjects should not have responded selectively to words having differing connotations for aggression.

One other source of ambiguity in the data should also be mentioned. The word "tear" obviously has a nonaggressive meaning, especially when presented visually to a subject so that it can be read as a noun describing the product of the lachrymal glands. Furthermore, the word "reduce" can have a violent meaning, as well as a nonviolent one, as in the sentence "The bombers reduced the village to rubble." We might assume that if either or both of these words elicited reactions in some subjects other than those expected from the pretesting, such reactions would result in fairly large variances in the aggressive reactions to the words. The variances in the intensity of aggression elicited by each of these words were compared to the variances associated with the remaining seven words, and in only one case did the difference in variances approach statistical significance: the difference between the variance of aggression intensity elicited by "tear" *vs.* that of aggression elicited by "annihilate" (1.55 *vs.* 2.14). The difference, however, was not significant ( $t = 1.41$ ,  $df = 22$ ,  $p < .10$ ).

The findings of the present study corroborate and extend the earlier ones of Geen and Stonner (8). Receiving reinforcement for aggressing leads to more intense levels of aggressiveness and, in addition, produces heightened aggressive reactions to stimuli that have relevance for violence. Furthermore, as association with violence of stimuli in a series decreases, so too does the degree of aggression elicited. The same is less true of subjects who are not reinforced for aggressing.

The results of the present study are consistent with Berkowitz's (4) viewpoint that aggressive behavior may be elicited by stimuli that have acquired strong aggressiveness associations through previous conditioning. Words such as "annihilate" and "kill" are probably usually encountered in situations containing either real or portrayed violence and may thereby become conditioned stimuli for feelings of aggressiveness originally elicited by violence. Our findings suggest that the greater is the individual's degree of aggressiveness habit strength, the more readily does the verbal conditioned stimulus elicit the aggressive conditioned response.

The results of this study also have implications for the trait-state problem in personality research. Recent discussions of personality theory have centered on the question of whether humans possess fixed and invariant characteristics that produce transsituational similarities in behavior, or whether behavior is mainly determined by variables present in a given situation (e.g., 13, 14).

Some writers (e.g., 1, 5) have suggested that behavior is the result of an interaction between "trait," or invariant, characteristics and "state," or situational variables. A trait may, for example, be considered a characteristic way of responding to situational stimuli or a determinant of one's threshold for responding to situational variation (6, 15). If we assume that reinforcement for aggression produces an "aggressive personality," we must conclude that behavior associated with that personality cannot be considered generally aggressive. Instead, it would appear that an aggressiveness trait interacts with situational stimuli: the greater the trait aggressiveness, the more powerfully will the person react to situational states that elicit aggression. However, when the situation does not elicit aggression, the "aggressive" personality behaves no more violently than the "nonaggressive."

## REFERENCES

1. ALKER, H. A. Is personality situationally specific of intrapsychically constant? *J. Personal.*, 1972, **40**, 1-16.
2. BANDURA, A., & WALTERS, R. H. *Social Learning and Personality Development*. New York: Holt, Rinehart & Winston, 1963.
3. BERKOWITZ, L. The concept of aggressive drive: Some additional considerations. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 2). New York: Academic Press, 1965.
4. ———. The contagion of violence: An S-R mediational analysis of some effects of observed aggression. In W. J. Arnold & M. M. Page (Eds.), *Nebraska Symposium on Motivation*. Lincoln: Univ. Nebraska Press, 1970.
5. ENDLER, N. S., & HUNT, J. McV. Sources of behavioral variance as measured by the S-R Inventory of Anxiousness. *Psychol. Bull.*, 1966, **65**, 336-346.
6. EYSENCK, H. J. *The Biological Basis of Personality*. Springfield, Ill.: Thomas, 1967.
7. GEEN, R. G., & PIGG, R. Acquisition of an aggressive response and its generalization to verbal behavior. *J. Personal. & Soc. Psychol.*, 1970, **15**, 165-170.
8. GEEN, R. G., & STONNER, D. Effects of aggressiveness habit strength on behavior in the presence of aggression-related stimuli. *J. Personal. & Soc. Psychol.*, 1971, **17**, 149-153.
9. GELLERMANN, S. The effects of experimentally induced aggression and inhibition on word-association sequences. Unpublished Doctoral dissertation, University of Pennsylvania, Philadelphia, 1956.
10. GRANT, D. A. Analysis of variance tests in the analysis and comparison of curves. *Psychol. Bull.*, 1956, **53**, 141-154.
11. GUILFORD, J. P. *Psychometric Methods*. New York: McGraw-Hill, 1954.
12. MILLER, N. Liberalization of basic S-R concepts: Extensions to conflict behavior, motivation, and social learning. In S. Koch (Ed.), *Psychology: A Study of a Science* (Vol. 2). New York: McGraw-Hill, 1959.
13. MISCHEL, W. *Personality and Assessment*. New York: Wiley, 1968.
14. ———. Continuity and change in personality. *Amer. Psychol.*, 1969, **24**, 1012-1018.
15. SPIELBERGER, C., Ed. *Anxiety and Behavior*. New York: Academic Press, 1966.

Department of Psychology  
University of Missouri-Columbia  
Columbia, Missouri 65201

## SO-CALLED "NERVOUS HABITS"\*

*School of Cultural and Community Studies, The University of Sussex, England*

---

D. G. WILLIAMS

---

### SUMMARY

The so-called "nervous habits" of 198 students were assessed by time-sampling 22 movement categories. Concentration of manipulatory behavior in the head and hand areas suggested that the stimulation generated by these activities is of some importance. Compared to men, women showed less facial manipulation and less leg and foot movement, but compensated, particularly by more nail-biting and manicure. Category scores did not have significant linear correlations with either Extraversion or Neuroticism scores on the Eysenck Personality Inventory (EPI). This may be due to idiosyncrasies in movement patterns.

---

### A. INTRODUCTION

"Nervous habit" is a loosely defined term popularly applied to such behavior as nail-biting, nonnutritive sucking, thumb-twiddling, various forms of self-scratching, and general restlessness (10, 11, 15). Koch (8) and Krout (10, 11, 12) proposed an "energy" or tension discharge function for nervous habits, and some other authors appear to accept this theory implicitly. However, studies of these behaviors in children (8, 15, 19, 21) have generally failed to find significant relationships with indices of emotionality or tension. Jones (6, 7) did find significantly more movements in adolescent subjects while they did mental arithmetic or suffered inhibition of micturition, but he did not correlate movement scores with personality measures. Sainsbury (18) found more arm muscle activity in a mixed group of adult psychotics and neurotics compared to healthy controls, and patients' preleucotomy scores were significantly correlated with clinical ratings of anxiety and tension. However, Sainsbury's electromyograph technique produced a single activity score per subject, and it was not shown that component arm and hand movements were equally related to a single factor of anxiety or tension. While Olson (15) believed nervous habits are an expression of a single factor, Koch

---

\* Received in the Editorial Office on November 6, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

(8) could find no evidence for this, and subsequent factor analysis (9) indicated that mannerisms had scattered loadings on nine factors. The results of Mehrabian and Williams (14) suggest at least that movements may be differentially related to neuroticism, and some may indicate relaxation during social interaction. Possible relationships with other personality variables, such as field dependence, are indicated by Freedman *et al.* (4).

The present study (a) tests the conventional hypothesis that the various types of "nervous" habits observed in the solitary individual are associated with high levels of neuroticism; (b) examines the "energy discharge" theory which in its simplest form predicts that, once sex differences are accounted for, activity will not be concentrated in a few body areas compared to equally accessible areas; and (c) examines and explains sex differences in frequency of different types of nervous habits. Sex differences in habit have been noted in children (8, 15, 19) and students (12, 13, 23), but have so far attracted little comment.

## B. METHOD

### 1. Category System

A system of 22 behavior categories was developed from previous systems (6, 8, 15) with more attempt to avoid loose conglomerates of diverse movement patterns. *Oral* included all rubbing of the mouth or teeth by one or both hands. Similarly, *Nasal*, *Aural*, *Ocular*, and *Hirsutal* involved manipulation of the nose, ears, eyes and orbits, and head and hair, respectively. *Facial* was any manipulation of the face, chin, or dewlap not designated Oral, Nasal, Aural, Ocular, or Hirsutal. *Biting* involved biting or sucking fingers or nails; *Manicure*, picking at fingers and manicuring nails; *Digital*, rubbing fingers together or rubbing hands; *Fidget*, manipulation of clothing and nonbody objects. *Lick* was biting or licking the lips; *Grimace*, pulling faces, frowning, smiling; *Eyeclose*, closing both eyes except blinking. *Respiratory* included whistling, yawning, sniffing, short coughs; *Gaze* was movement of the head, except *Caputal* which was head movement with eyes remaining focused on the same point. *Leg* and *Pedal* involved leg and foot movement, respectively, and *Body* was rotation of the body about the vertical or horizontal axes and shifting buttocks on seat. *Arm* included gesturing and hand movement from one body region to another; *Irritational*, manipulation of any part of the body except head and hands; and *Talk*, talking or whispering to self.

### 2. Procedure

Subjects were viewed through a one-way screen as they sat alone trying to solve a difficult "trick" problem that they believed 60 percent of students



could solve within 40 minutes. *S* sat in an upright armchair half-facing the screen about eight feet from *E*. Movements were time-sampled by the intermittent method (16) for 120 five-second observation periods, with 10-second intervals between periods during which one check mark was entered on a prepared record sheet for each category observed for all or any part of the preceding period. Timing was regulated by prerecorded tape. Observations were suspended if *S* attempted a solution or behaved in any way that made some categories unobservable. Observations were abandoned if suspensions totalled more than four minutes. After 40 minutes *E* re-entered the test room, and in subsequent debriefing any suspicious *S* who could identify the means of observation was rejected. Finally each *S* completed forms A and B of the Eysenck Personality Inventory (3) which scores Neuroticism (N) and Extraversion (E).

### 3. Subjects

The sample of 198 university students did not differ significantly from 98 rejectees and pilot subjects in EPI N or E scores. The 97 male *Ss* were older than the 101 female *Ss* (Mann-Whitney  $z = -3.12$ ,  $p < .01$ ) although difference in mean age amounted to only 9.5 months (male 20.5 years, *SD* 2.0; female 19.7 years, *SD* 1.3). Male EPI N scores averaged 22.7 (*SD* 9.6), and female were 15.9 (*SD* 8.0), this difference being significant (Mann-Whitney  $z = 2.58$ ,  $p < .01$ ). Male E scores averaged 26.2 (*SD* 7.5), female 26.5 (*SD* 7.7), an insignificant difference.

### C. RESULTS

Gaze obtained the highest mean score of all categories ( $\bar{X} = 29.3$ , *SD* = 12.5). Comparable means and *SDs*, respectively, for other categories were as follows: Facial (27.0, 16.6); Oral (21.0, 13.6); Caputal (11.0, 6.8); Arm (10.9, 6.1); Irritational (8.2, 5.8); Eyeclose (14.6, 20.8); Pedal (10.8, 12.2); Digital (10.7, 10.0); Nasal (8.7, 7.8); Ocular (7.7, 7.6); Hirsutal (6.4, 11.1); Body (5.9, 4.6); Leg (5.1, 7.5); Lick (4.9, 5.4); Grimace (4.3, 4.6); Manicure (4.9, 9.6); Fidget (4.5, 7.3); Bite (4.3, 9.6); Talk (3.0, 6.7); Aural (2.4, 2.6); Respiratory (2.1, 2.4). All category distributions except Gaze were significantly positively skewed (each  $p < .01$ ).

Rulon split-half reliabilities were satisfactorily high and consistent with earlier studies, the lowest coefficient being Respiratory (.64), the highest Eyeclose (.99), and 13 categories being over .90. It is debatable whether such figures are particularly meaningful (1). Accurate observation is more realistically indicated by interpretable sex differences and findings of lower Ocular

scores amongst the 28 Ss who wore corrective lenses ( $\chi^2 = 5.71$ ,  $df = 1$ ,  $p < .05$ ).

A Sign Test (20) comparison of Irritational scores with categories of head and hand manipulation produced the following  $z$  scores, where a negative sign indicates higher Irritational: Irritational *v.* Digital .9, *v.* Oral 8.9 ( $p < .01$ ), *v.* Facial 12.2 ( $p < .01$ ), *v.* Nasal — 1.1, *v.* Aural — 11.6 ( $p < .01$ ), *v.* Hirsutal — 7.7 ( $p < .01$ ), *v.* Ocular — 2.8 ( $p < .01$ ). Thus Oral and Facial scores were significantly higher than Irritational scores; Digital scores were slightly so. This pattern held even if male and female results were analyzed separately.

Sex differences were analyzed by Mann-Whitney  $U$  test (20) and checked by chi square. Men showed significantly higher levels of Oral (Mann-Whitney  $z = -4.13$ ,  $p < .01$ ), Facial ( $z = -3.66$ ,  $p < .01$ ), Nasal ( $z = -3.37$ ,  $p < .01$ ), and Ocular ( $z = -2.07$ ,  $p < .05$ ) manipulation; more Eyeclose ( $z = -3.01$ ,  $p < .01$ ); and more Leg ( $z = -3.92$ ,  $p < .01$ ) and Pedal ( $z = -2.19$ ),  $p < .05$ ) activity. Women showed significantly more Bite ( $z = 2.43$ ,  $p < .01$ ), Manicure ( $z = 5.01$ ,  $p < .01$ ), Fidget ( $z = 2.35$ ,  $p < .05$ ), Irritational ( $z = 2.59$ ,  $p < .01$ ), Arm ( $z = 2.34$ ,  $p < .05$ ), and Hirsutal ( $z = 2.33$ ,  $p < .05$ ) behavior. Sex differences were insignificant in Gaze ( $z = -.2$ ), Caputal (—1.1), Digital (1.6), Body (—1.5), Lick (—1.0), Grimace (—1.5), Talk (—4), Aural (.8), and Respiratory (—7) categories.

Category distributions were normalized, where possible, and product-moment correlations with E and N calculated. Biserial correlations were calculated for categories with nontransformable distributions. No category had a significant correlation with E, nor did any significant correlation emerge from partial correlation extracting sex and N variance. The only categories significantly correlated with N were Oral ( $r = -.16$ ,  $p < .05$ ), Hirsutal ( $r = .17$ ,  $p < .05$ ), and Fidget ( $r_{bis} = .18$ ,  $p < .05$ ), but extraction of sex variance reduced the Oral correlation to insignificance ( $r = -.12$ ). Thus only two out of 22 categories appeared significantly positively correlated with N; the variance attributable to N is trivial, and at least one of these correlations would be expected through chance effects.

#### D. DISCUSSION

If "nervous habits" are performed merely to discharge tension or "surplus energy," then high levels of activity would be expected in the Irritational category because of the comparatively large body surface area involved. Concentration of manipulation in head and hand regions suggests rather that activities are performed, at least partly, because of the feedback stimulation

that they provide. The head and hands, not being covered with clothing, are presumably more sensitive to manipulation than other areas. Secondly, face, mouth, and hands have proportionately greater cortical representation (17). This may explain why Aural, Hirsutal, Nasal, and Ocular activities are low compared to Oral and Facial, although such features as ease of performance probably play some part.

Feedback stimulation may be useful because it is pleasant or satisfying in some way, because it is selectively attended to in preference to more stressful events, or because it swamps or suppresses stressful stimuli, as in dental patients who raise their pain thresholds by tightly gripping their chairs (5). Jones's (6, 7) studies suggest that different stressors require different combinations of movements to provide appropriate feedback.

Although women spent more time playing with and adjusting their hair, they spent less time rubbing the mouth, face, nose, and eyes, presumably for cosmetic reasons. Men seemed more physically active, particularly showing more leg and foot movement. The wearing of dresses and other characteristic sartorial features may inhibit much of this activity in women. Movements involved in nail-biting produced higher Arm scores in women. Although men were more active, longer periods spent sitting comfortably manipulating the face led to more relaxed eyeclosing. Female levels of manicuring and nail-biting may be partly due to greater concern for appearance, but otherwise would seem to be compensation for lessened activity in other areas, as would the higher levels of hand rubbing, manipulation of nonbody objects, and body scratching. This suggests that although the sexes differ in their patterns of movement, they differ less in their total levels of activity. Cultural changes may lead to greater similarity in movement pattern between the sexes.

Freedman *et al.* (4) found that behavior corresponding to the Digital category correlated with field dependence in a small number of female Ss engaged in social interaction. Field dependence is related to E (2), but no "nervous habit" category in this study of solitary individuals was correlated significantly with E. However, failure to find relationships between category scores and N was more surprising. Subjects may have been too homogeneous to allow relationships to emerge effectively, or personality questionnaires emphasising long-term, average characteristics may be insufficiently close measures of actual disturbance in the test situation. The most likely explanation is that nervous habit patterns are idiosyncratic, and this obscures any relationship with neuroticism except under special circumstances. Sainsbury (18) found that each of his patients had individual mannerisms which regularly recurred in similar situations. In comparing patients before and after leu-

cotomy, each acts as his own control, and personal idiosyncracies are partially balanced. Similar controls occurred in Jones's experiments when subjects were studied under different stress levels. Finally, in comparing psychiatric and normal groups, Sainsbury's EMG technique masks idiosyncrasy by producing a single activity measure associated with a variety of arm and hand movements.

Although Vernallis' (22) study of teeth grinding indicates that more unusual habits may be related to anxiety, it would seem that more commonly observed habits cannot be taken on their own as symptomatic of disturbance. Some very stable subjects will be amongst the highest scorers on any particular category.

### REFERENCES

1. ARRINGTON, R. E. Time-sampling in studies of social behavior. *Psychol. Bull.*, 1943, **40**, 81-124.
2. EVANS, F. J. Field dependence and the Maudsley Personality Inventory. *Percept. & Motor Skills*, 1967, **24**, 526.
3. EYSENCK, H. J., & EYSENCK, S. B. G. *Manual of the Eysenck Personality Inventory*. London: Univ. London Press, 1964.
4. FREEDMAN, N., O'HANLON, J., OLTMAN, P., & WITKIN, H. A. The imprint of psychological differentiation on kinetic behavior in varying communicative contexts. *J. Abn. Psychol.*, 1972, **79**, 239-258.
5. ISCHLONDSKY, N. E. *Brain and Behaviour*. London: Kimpton, 1949.
6. JONES, M. R. Studies in "nervous" movements: I. The effect of mental arithmetic on the frequency and patterning of movements. *J. Gen. Psychol.*, 1943, **29**, 47-62.
7. ———. Studies in "nervous" movements: II. The effect of inhibition of micturition on the frequency and patterning of movements. *J. Gen. Psychol.*, 1943, **29**, 303-312.
8. KOCH, H. L. An analysis of certain forms of so-called "nervous habits" in young children. *J. Genet. Psychol.*, 1935, **46**, 139-169.
9. ———. A factor analysis of some measures of the behavior of preschool children. *J. Gen. Psychol.*, 1942, **27**, 257-287.
10. KROUT, M. H. A preliminary note on some obscure symbolic muscular responses of diagnostic value in the study of normal subjects. *Amer. J. Psychiat.*, 1931, **11**, 29-71.
11. ———. Autistic gestures: An experimental study in symbolic movement. *Psychol. Monog.*, 1935, **46**, No. 208.
12. ———. An experimental attempt to produce unconscious manual symbolic movements. *J. Gen. Psychol.*, 1954, **51**, 93-120.
13. ———. An experimental attempt to determine the significance of unconscious manual symbolic movements. *J. Gen. Psychol.*, 1954, **51**, 121-152.
14. MEHRABIAN, A., & WILLIAMS, M. Nonverbal concomitants of perceived and intended persuasiveness. *J. Personal. & Soc. Psychol.*, 1969, **13**, 37-58.
15. OLSON, W. C. *The Measurement of Nervous Habits in Normal Children*. Minneapolis: Univ. Minneapolis Press, 1929.
16. OLSON, W. C., & CUNNINGHAM, E. N. Time-sampling techniques. *Child Devel.*, 1934, **5**, 41-58.
17. PENFIELD, W. G., & RASMUSSEN, T. *The Cerebral Cortex of Man: A Clinical Study of Localization of Function*. New York: Macmillan, 1950.



18. SAINSBURY, P. The measurement and description of spontaneous movements before and after leucotomy. *J. Ment. Sci.*, 1954, **100**, 731-741.
19. SEHAM, M., & BOARDMAN, D. V. A study of motor automatisms. *Arch. Neurol. Psychiat.*, 1934, **32**, 154-173.
20. SIEGEL, S. *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill, 1956.
21. TRYON, A. F. Thumb-sucking and manifest anxiety: A note. *Child Devel.*, 1968, **39**, 1158-1163.
22. VERNALLIS, F. Teeth-grinding: Some relationships to anxiety, hostility and hyperactivity. *J. Clin. Psychol.*, 1955, **11**, 389-391.
23. YOUNG, F. M. The incidence of nervous habits observed in college students. *J. Personal.*, 1947, **15**, 309-320.

*School of Cultural and Community Studies*  
*The University of Sussex*  
*Arts Building*  
*Falmer, Brighton BN1 9QN*  
*England*



## CREATIVITY IN RURAL, URBAN, AND INDIAN CHILDREN\*

*The Center for Teaching and Learning, The University of North Dakota*

---

JOHN D. WILLIAMS, JOHANNA TEUBNER,<sup>1</sup> AND STEVEN D. HARLOW

---

### SUMMARY

Five groups of grade school children ( $N = 237$ ) were selected and administered the *Torrance Tests of Creative Thinking*. The five groups included urban-middle income children ( $N = 47$ ), urban-lower income children ( $N = 49$ ), rural children ( $N = 72$ ), Indian-lower income children ( $N = 54$ ), and Indian-impooverished children ( $N = 15$ ). Significant differences were found on all three measures of verbal creativity, with the verbal flexibility and verbal originality measures significant at the .001 level, and the verbal fluency measure significant at the .005 level. On all three measures, the groups scored from high to low in the following order: rural, urban-middle income, urban-lower income, Indian-lower income, and Indian-impooverished. None of the four measures of figural creativity showed significant differences.

---

### A. INTRODUCTION

A great deal of effort has gone into the study of creativity in the past two decades. Much of that effort has been directed toward a better understanding of the construct of creativity, and of the various dimensions that have been posited to constitute or contribute to the operation and concept of creativity. Concurrently, many experimental and quasi-experimental studies have operationalized one or another of the concepts of creativity in relation to an experimental variable. To a limited extent, some cross-cultural studies of creativity have been completed. Torrance (4) investigated perceived pressure against divergent thinking among three groups of New World preadolescent children (Twin Cities and International Falls, Minnesota; and Puerto Rico) and a similar age group of Old World children (England, France, Turkey, and Greece). More pressure was perceived against divergent thinking in the New World groups. Doyle (1) found no differences between black and white children on measures of verbal fluency and creative talent, although there was

---

\* Received in the Editorial Office on November 10, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Now at Lewis-Clark State College, Lewiston, Idaho.

some tendency for the blacks to score higher on the latter measure. Using the *Minnesota Tests of Creative Thinking*, Irons (2) found, with a sample of children from Northeast Texas, that urban children were generally more creative than rural children. A study (4) was completed in New Mexico using the *Minnesota Tests of Creative Thinking* comparing Anglo, Indian, and Spanish-surnamed children. The Anglos were found to be significantly more creative than the other two groups; the Spanish-surnamed children outperformed the Indian children, though the difference between these two groups was nonsignificant.

Torrance (6) has investigated any possible bias for or against disadvantaged groups in regard to the *Torrance (5) Tests of Creative Thinking*. He reported that only slight differences were found between children from black and white groups; similarly, little difference was evidenced between middle and low income groups.

## B. METHOD

### 1. Subjects

The subjects for the present study were selected from fourth grade classrooms in several North Dakota schools. Five different groups were defined. Two groups were chosen as being representative middle income or lower income urban children; two classrooms were selected for each of the income levels. Four classrooms were selected in three different rural communities and were considered representative of rural children. Four classrooms were selected from two Indian reservation areas. Three of the four selected classrooms were considered typical, in a socioeconomic sense, of the reservation areas; these three classrooms would be considered to be in lower income areas. A fourth Indian classroom was selected from an area that can be described as being economically impoverished. Thus, five groups were defined for this study: urban-middle income ( $N = 47$ ), urban-lower income ( $N = 49$ ), rural ( $N = 72$ ), Indian-lower income ( $N = 54$ ), and Indian-impoverished ( $N = 15$ ). Each subject was administered the *Torrance Tests of Creative Thinking*, including the verbal and figural sections.

### 2. Procedure

The Ask-and-Guess portion of the test has three activities which are all based on a drawing. The first activity requires the student to write all the questions that he can think of that would be helpful for understanding the picture. The second activity asks the student to list as many possible causes as he can for the action shown in the drawing. The third activity asks the



student to list as many possibilities as he can concerning what might happen as a result of the activity in the picture. Product Improvement, the fourth activity, asks the student to list the cleverest, most interesting, and unusual ways for changing a toy elephant so that children will have more fun playing with the elephant. Unusual Uses of Cardboard Boxes, the fifth activity, asks the student to list as many interesting and unusual uses of cardboard boxes as he can. The sixth activity, Unusual Questions About Cardboard Boxes, asks the student to think of questions about aspects of cardboard boxes which people do not usually think about. Just Suppose, the seventh activity, gives the student an improbable situation. The child has to suppose that it has happened, and as a consequence of this he is to think of all of the other things that would happen as a result of this improbable situation. The sixth activity is scored for verbal fluency; the remaining six verbal activities are scored for verbal fluency, verbal flexibility, and verbal originality.

The first activity of the figural test is Picture Construction. The child is given a piece of colored paper in the form of a curved shape: the child is to stick a colored shape wherever he wishes to make a picture he has in mind; then he is asked to add lines with his pencil or crayon to make a picture. This activity gives a score for originality and elaboration. Picture Completion, the second figural activity, asks the child to add lines to incomplete figures, thereby sketching some interesting objects or pictures, and then to make an interesting title for each drawing. The third activity, Lines, is a task that asks the child to make as many objects or pictures as he can from the pairs of straight lines. The pairs of straight lines should be the main part of whatever is made. With pencil or crayon the child can add lines to the pairs of lines to complete the picture. The last two tasks are scored for figural originality, figural elaboration, figural fluency, and figural flexibility.

### C. RESULTS

From Table 1 it can be seen that significant differences are found on all three measures of verbal creativity, but no significant differences were found for figural creativity. The verbal flexibility and verbal originality measures both show significance at the .001 level, while the verbal fluency measure is significant at the .005 level. Further inspection of the three significant verbal measures yields a systematic result: the rural group scored highest on all three measures, the urban-middle income group scored second highest on all three measures, the urban-lower income group scored third on all three measures, and the two Indian groups scored the lowest, with the Indian-impooverished group being the lower of the two. On the verbal flexibility mea-

TABLE 1  
MEANS AND *F* VALUES FOR THE TORRANCE TESTS OF CREATIVE THINKING

Variable	Urban-middle income	Urban-lower income	Rural	Indian-lower income	Indian- impoverished	<i>F</i>	<i>p</i>
Verbal fluency	39.36	38.89	40.69	36.94	32.47	4.29	.005
Verbal flexibility	49.56	48.06	51.74	44.49	32.33	8.62	.001
Verbal originality	44.36	43.89	44.44	41.84	37.00	5.39	.001
Figural fluency	42.23	41.85	39.10	40.41	40.00	1.52	N.S.
Figural flexibility	47.77	45.93	44.72	46.84	45.33	1.08	N.S.
Figural originality	46.81	43.89	44.72	42.33	42.33	1.13	N.S.
Figural elaboration	54.89	52.32	52.28	49.33	49.33	.96	N.S.

sure, the Indian-impoveryished group scored markedly lower than the other four groups.

The results for the figural creativity portion show no significant differences; with one notable exception, the results follow in approximately the same order the results for verbal creativity. That exception is the rural children; while they scored the highest on all verbal variables, they scored *lowest* on the measures of figural fluency and figural flexibility. The rural group scored second highest on figural originality, and third on figural elaboration.

Within the urban groups, the middle income group scored higher on all seven measures of creativity; the differences between these two groups are not marked on any one variable, however. Similarly, the Indian-lower income group scored higher on all seven measures of creativity when compared to the Indian-impoveryished group: the differences in means on the verbal measures all exceed four units; on the verbal flexibility measure, the difference is 12.16 units.

#### D. DISCUSSION

It does appear that, to some extent, socioeconomic status is related to success on the *Torrance Tests of Creative Thinking* in that, for the urban children, the middle income group scored higher on each of the seven measures of creativity; similarly, the Indian-lower income group scored higher than the Indian-impoveryished group on the same seven variables. The lower scores on these tests by the Indians when contrasted to non-Indians is in keeping with Mayhon's (3) findings. While several reasons might be posited for this difference, it might be pointed out that Indian children in general do not see Indian adults in such positions as teachers, administrators, or researchers, and they may tend to see such experiences as taking tests as being one of the "white man's games." Their improved status in the figural measures as compared to the verbal measures might also indicate to some degree a linguistic barrier.

The data presented herein would imply that rural children do not suffer the deficit in creativity that was shown in Iron's (2) study. The difference in the results of the present study with Iron's study may be to some degree a regional difference in the status of the farmer relative to urban groups. Because of the rural nature of North Dakota (the most rural state in the nation), being from a rural area is viewed by the rural dweller to some degree as a source of pride. Thus rural North Dakota children in general might have high self-concepts which, in turn, might be translated into enabling potential for expressing their creative abilities.

## REFERENCES

1. DOYLE, M. A. A comparative study of creativity in Negroes and Caucasians. *Child Study J.*, 1970, **1**, 25-27.
2. IRONS, J. L. Creative thinking about abilities of rural and urban elementary school students. *Diss. Abst.*, 1968, **28**, 2897-A.
3. MAYHON, W. G. The relationship of creativity to achievement and other student variables. *Diss. Abst.*, 1966, **27**, 1713-A.
4. TORRANCE, E. P. Role of Evaluation in Creative Thinking. Minneapolis: Bur. Educ. Res., Univ. Minnesota, 1964.
5. ———. Torrance Tests of Creative Thinking. Princeton, N. J.: Personnel Press, 1966.
6. ———. Are the Torrance Tests of Creative Thinking biased against or in favor of "disadvantaged groups." *Gifted Child Quart.*, 1971, **15**, 75-80.

*The Center for Teaching and Learning*

*The University of North Dakota*

*Grand Forks, North Dakota 58201*



## BRAINSTORMING IN LARGE GROUPS AS A FACILITATOR OF CHILDREN'S CREATIVE RESPONSES\* 1, 2

*Department of Psychology, California State University, San Francisco*

LYLE J. BUCHANAN, JR., AND HENRY CLAY LINDGREN<sup>3</sup>

### SUMMARY

Six groups of fourth graders ( $N = 122$ ) from three private schools seated in formal classroom settings were given an unusual-uses problem, and for a 10-minute period they either brainstormed as a total group or brainstormed individually. In the second phase of the experiment, all groups brainstormed individually for five minutes. In two schools (four groups), those who had engaged in group brainstorming in the first phase made significantly more responses and demonstrated a higher level of creativity than those who had brainstormed individually in both phases. Differences for the third school were not significant. Results for these two schools with their traditional classrooms are consistent with the results of experiments with adults showing that brainstorming in small, informal groups tends to facilitate creativity in subsequent problem-solving sessions. Additional research is needed to determine the extent to which such factors as general arousal, normative effect, imitation, or reinforcement contribute to the enhancement of creative responses in group brainstorming situations.

### A. INTRODUCTION

A number of experimenters have noted that brainstorming in small groups has a facilitative effect on creativity in subsequent individual brainstorming (1, 2, 3, 4). Lindgren and Lindgren (4), for example, found that university students who first brainstormed alone in writing cartoon captions, then brain-

\* Received in the Editorial Office on November 14, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Based on the first author's thesis, submitted in partial fulfillment of the requirements for the M. S. degree in psychology, California State University, San Francisco, 1973.

<sup>2</sup> The authors express appreciation for the assistance of Domenico Giglio, who served as one of the experimenters, and for the help of Dr. Rolfe LaForge, Department of Psychology, California State University, San Francisco, who advised on statistical treatment of the data.

<sup>3</sup> Address reprint requests to the second author at the address shown at the end of this article.

stormed in small groups assigned the same type of task, and finally brainstormed alone, produced a larger number of responses and more creative responses in the third phase of the experiment than they had in the initial phase. Inasmuch as a control group that had not engaged in group brainstorming showed no such improvement, the enhancement of creativeness in the experimental group's third phase was attributed to the experimental treatment they had experienced in the second phase.

The present study was undertaken to determine whether (a) a similar facilitation of creativity through group brainstorming could be achieved with elementary school children, and (b) group brainstorming in formal classroom settings would be effective in stimulating creativity in subsequent individual sessions.

### B. METHOD

Subjects participating in the study were 122 fourth graders of both sexes attending three private schools having largely middle-class enrollments. At each school, subjects were randomly assigned to experimental or control treatments and seated in different classrooms. The two treatments were run simultaneously, with one experimenter randomly taking the experimental group, and the other, the control group. Tasks were six items utilizing the technique of Unusual Uses from Torrance's Test of Imagination, Form D (5), and were presented in randomized order in the two phases of the experiment.

Subjects in the experimental group were seated in formal, straight rows, all facing the experimenter, in traditional classroom style. Subjects were asked to brainstorm by thinking of as many unusual, interesting, and clever uses for the stimulus object (shoe box, paper clip, or whatever), to say them aloud so that everyone could hear them, and to write down their own responses, as well as others'.

This first phase of the experiment was concluded at the end of 10 minutes, whereupon subjects were given another item as a stimulus object and told to record on a second sheet of paper as many unusual, interesting, and clever uses for the second stimulus object as they could. During this individual brainstorming phase, which lasted five minutes, subjects did not speak.

Control groups sat in similar classrooms and brainstormed individually and silently for the same periods of time for both phases of the experiment (10 minutes for phase one and five minutes for phase two).

Creativity was measured in terms of the number of responses each subject produced during the second phase (individual brainstorming), and the level

of creativity expressed in whatever responses he did produce. Level-of-creativity scores were obtained by summing the ratings assigned by three judges who operated independently. Judges were instructed to rate each set of responses in terms of their overall creativity—that is, the extent to which the responses could be characterized as original, clever, and interesting, as well as appropriate to the stimulus item in question. Response sheets completed by experimental and control subjects were first coded in a way that could not be recognized by the judges and then mixed together. Then they were sorted into batches of 16 or 17 sheets each and were kept separate for each school. The sheets in each batch were then rated on a seven-point scale in such a way that judgments were forced into an approximation of a crude normal curve. One response sheet in each batch was assigned the rating of "7" (most creative), and one sheet the rating of "1" (least creative). Two sheets were each given a rating of "6," and two were given a rating of "2," and so on. After a judge had completed the entire lot of 122 response sheets and recorded his scores, the response sheets of each school were reshuffled and randomly sorted into batches of 16 for the next judge. Interjudge correlations (Cronbach's alpha) among the three judges were .83, .82, and .88, which compare favorably with those obtained by Lindgren and Lindgren (3, 4), who used the same method in assessing the level of creativity of cartoon captions.

### C. RESULTS

Stepwise multiple regression analyses revealed that there was no significant relationship between subjects' sex, age, and grade-placement scores on the SRA Achievement Test, Form D (grades 2-4), on the one hand, and number of responses, or level-of-creativity score, on the other.

Stepwise multiple regression analyses did show, however, that the experimental groups in Schools 2 and 3 made significantly more responses ( $p < .01$ ) and demonstrated a higher level-of-creativity ( $p < .01$ ) than did the control groups. There were no significant differences between the experimental and control groups at School 1. Table 1 gives the means and standard deviations for the number of responses, and Table 2, for the level-of-creativity scores.

### D. DISCUSSION

Results for two out of the three schools indicate that group brainstorming apparently has a subsequent facilitative effect on creativity for elementary children, just as it had for adults in previous studies, and formal, traditional classrooms apparently do not inhibit the effect.

The anomalous results from School 1 may have been due to unanticipated

TABLE 1  
NUMBER OF RESPONSES FOR PHASE ONE AND PHASE TWO OF EXPERIMENT

Samples	Experimental groups			Control groups			
	Phase 1—group brainstorming		Phase 2—individual brainstorming	Phase 1—individual brainstorming			Phase 2—individual brainstorming
	<i>N</i>	Number <sup>a</sup>	<i>M</i> <i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>M</i> <i>SD</i>
School 1	27	11	5.8      2.8	24	5.2	3.1	5.3      3.3
School 2	16	30	17.8      5.8	22	4.4	3.5	4.5      3.8
School 3	19	35	20.8      5.9	14	9.1	5.4	10.1      5.7

<sup>a</sup> Total group responses.



TABLE 2  
 MEANS AND STANDARD DEVIATIONS OF PHASE TWO OF LEVEL-OF-CREATIVITY SCORES

Samples	Experimental groups			Control groups		
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
School 1	27	12.2	2.9	24	11.7	3.3
School 2	16	15.1	2.5	22	9.3	3.1
School 3	19	13.5	3.7	14	9.8	3.3

interference on the part of a teacher who insisted on remaining in the classroom during the experimental treatment in order to "help." He supplemented the experimenter's directions by ordering the children to cooperate, and later told the experimenter that he had done so because the school made no use of group discussion techniques, and the children, as a consequence, had had no experience with the method. Whatever the cause, the children in that school did not have a genuine group brainstorming experience. Table 1 shows that during the first phase of the experimental treatment in School 1, they produced only 11 responses, or an average of .41 per child, in contrast with averages of 1.88 for School 2 and 1.85 for School 3.

Another contributing factor to the poor performance of School 1 experimental subjects may have been the size of the group, which was half again larger than the corresponding groups in the other two schools. In other words, it may be that the size of the group is a limiting factor to the facilitating effects of group brainstorming on subsequent creativity.

If group brainstorming facilitates creativity in subsequent problem-solving sessions in the normal classroom situation, as the results from at least two of the schools suggest, further research should be directed to determining which aspect of the group brainstorming experience produces the facilitation effect. One possibility would appear to be *general arousal*—the result of participating in an active discussion. Competitiveness may also be a side-effect here. Enhanced performance may also be the result of the task becoming more salient. Such an explanation would depend on the operation of *group norms*: an individual who is not particularly interested in the kind of task set by the experimenter finds, as he observes other group members participating eagerly, that the task is more interesting than he had thought earlier. Still another explanation is *imitation*—the tendency of slower individuals to use more able group members as models. Members who participate in group brainstorming also find *reinforcement* in group approval, and even less active group members may find their subsequent performance enhanced as a vicarious gain from having observed the reinforced performance of others.

Each of these explanations appears to have merit, and their relative contribution to subsequent response enhancement can be determined only by a series of carefully controlled experiments.

## REFERENCES

1. DUNNETTE, M. D., CAMPBELL, J., & JAASAD, K. The effect of group participation on brainstorming effectiveness for two industrial samples. *J. Appl. Psychol.*, 1963, **47**, 30-37.
2. LINDGREN, H. C. Brainstorming and the facilitation of creativity expressed in drawing. *Psychol. Rep.*, 1967, **24**, 350.
3. LINDGREN, H. C., & LINDGREN, F. Brainstorming and orneriness as facilitators of creativity. *Psychol. Rep.*, 1965, **16**, 577-583.
4. LINDGREN, H. C., & LINDGREN, F. Creativity, brainstorming, and orneriness: A cross-cultural study. *J. Soc. Psychol.*, 1965, **67**, 23-30.
5. TORRANCE, E. P. *Rewarding Creative Behavior*. Englewood Cliffs, N. J.: Prentice Hall, 1965.

120 Lansdale Avenue  
San Francisco, California 94127

## DRUG USAGE, PERSONALITY, ATTITUDINAL, AND BEHAVIORAL CORRELATES OF DRIVING BEHAVIOR\*<sup>1</sup>

*University of California, Los Angeles*

---

KAY JAMISON AND WILLIAM H. MCGLOTHLIN

---

### SUMMARY

Subjects ( $N = 164$ ) were divided into four groups on the basis of their driving records: no accidents or moving violations, two or more violations, one or more accidents, and no accidents. They were compared on various personality, attitudinal, drug usage, and biographical variables. The less safe drivers scored significantly higher on a sensation-seeking scale; were more likely to have experimented with such drugs as marijuana, hashish, amphetamines, LSD, etc.; attended religious services less often; and were less likely to own their residences. There were also differences in political attitudes and life-style preferences among the groups.

---

### A. INTRODUCTION

The multifaceted problem of predicting driving behavior has been studied for many years by investigators with varying methods of approach. A sizeable proportion of the research studies has centered on motor and perceptual capacities of drivers and on such background variables of drivers as sex, age, and amount of driving experience. Increasingly, however, there have been attempts to determine what proportion of the variance in driving behavior can be accounted for by personality and attitudinal variables. The purpose of the study outlined in this paper was to investigate further some of the drug usage, personality, attitudinal, and biographical correlates of driving behavior in an attempt to replicate and expand on the findings of earlier work.

### B. PROCEDURE

#### 1. Sample

The data for this study were obtained from a larger sample of 247 subjects who had been given LSD as part of psychotherapy or under experimental con-

---

\* Received in the Editorial Office on November 15, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was supported by U.S.P.H.S. Grant MH-13484 (NIMH) and by a U.S.P.H.S. predoctoral research fellowship to the first author.

ditions several years ago (1955-1961). They were recently interviewed and given a battery of personality and attitude scales as part of an investigation into the long-term effects of LSD (13). The subsample used here comprised those persons, of the original 247, for whom driving records covering the past three years could be obtained from the California Department of Motor Vehicles ( $N = 164$ , or 66% of the larger sample). The final sample for this study was 67% male, mean age 40, considerably above the national norm on educational level (52% had a B.A. or higher degree); about 80% had been involved in psychotherapy at one time or another.

## 2. Inventories Used

The biographical and drug information was acquired by means of a personal interview, the attitudinal and personality data via a questionnaire left with the subjects to complete and return. The aphorism scale used was a modified version of a scale developed by McGlothlin *et al.* (14) in their study of the long-lasting effects of LSD on normals. Aphorisms relating to the importance of self-knowledge, overcoming egocentrism, mystical orientation, and a passive philosophy were judged on the positive end of the scale. Belief in paranormal phenomena was measured by a series of items tapping how much the subject believed in such phenomena as UFOs, astrology, reincarnation, ESP, etc. The Myers-Briggs Type Indicator (15), shortened form, was used to get a measure of preference for (a) intuition and ideas over conventional and factual approaches (the sensing-intuition subscale) and (b) a casual, spontaneous style of life as opposed to one that is orderly and systematic (the judgment-perception subscale). The shortened form of the Marlowe-Crowne Social Desirability Scale (7) consisted of items defined by behavior that is "culturally sanctioned and approved but improbable of occurrence." Other personality tests included Aas' Hypnotic Suggestibility Scale (1), a sensation-seeking scale developed by Zuckerman *et al.* (21), and a semantic differential scale. Scales to test attitudes about work, political issues, individual freedom, the legal system, etc. were also used, in addition to a political activity index developed by Woodward and Roper (20) and a modified version of the organizational activity index of Chapin (4). Items to measure alienation were taken from Srole (18).

## 3. Data Analysis

Subjects were divided into four groups on the basis of their three-year driving records: Group I—those who had had no accidents or moving violations ( $N = 52$ ); Group II—those who had had two or more moving violations ( $N = 61$ ); Group III—those who had had one or more accidents ( $N = 43$ );



and Group IV—those who had had no accidents ( $N = 121$ ). A frequency distribution program was run<sup>2</sup> which obtained means, standard deviations, and percentages for different variable categories. Chi squares or  $t$  tests were used to examine differences between Groups I and II and between Groups III and IV.

A multiple regression analysis was also done after obtaining a correlation matrix for 10 predictor variables (sex, owning *vs.* renting of residence, experimentation with marijuana and hashish, frequency of attendance at religious services, and scores on the aphorism, Myers-Briggs judgment-perception, sensation-seeking, relativity of sanity, and LSD cultism scales) and four criteria (dichotomous categories based on the four subject groupings).

### C. RESULTS

#### 1. *Biographical Data*

As Table 1 shows, there were virtually no differences among the groups on the basis of age, educational level, history of broken families, or judged happiness of parents' marriage. The socioeconomic status, percentage involved in psychotherapy, and marital stability of the no accidents or violations group was lower than that of the two or more violations group, but the opposite was true for the no accidents *vs.* one or more accidents group; none of these differences was statistically significant.

The percentage of first or only children was higher in both the two or more violations group and the one or more accidents group than in their respective comparison groups; although not significant, the differences were in the direction predicted by Adler's theory (2) and Schacter's empirical work (16). The two biographical variables that showed a significant difference (chi square,  $p < .01$ ) between the no accidents or violation group *vs.* the two or more violations group were sex (more males in the latter group) and owning *vs.* renting of residences (more owning of residences in former group).

#### 2. *Drug Usage Data*

The two or more violations group had a higher percentage of experimentation or usage in every drug category, except for sedatives, than did the no accidents or violations group (see Table 2). The differences approached significance for the usage of amphetamines [in line with Smart's (17) findings], opium, morphine, or cocaine; the strong hallucinogens other than LSD

<sup>2</sup> Computing assistance was obtained from the Health Sciences Computing Facility, U.C.L.A., sponsored by NIH Special Resources Grant RR-3 and from the Campus Computing Network.

TABLE 1  
BIOGRAPHICAL DATA

Variable	Group			
	I. No accidents or violations (N = 52)	II. Two or more violations (N = 61)	III. One or more accidents (N = 43)	IV. No accidents (N = 121)
Age (mean)	45.2	43.2	44.0	43.8
Sex (% male)	50.0*	78.7	69.8	64.5
Socioeconomic status (mean)	84.9	90.0	86.6	88.6
Education (mean) 1 = 8th grade or less; 8 = doctorate	5.7	5.7	5.5	5.8
Separated, divorced, or remarried (%)	50.0	40.9	46.4	44.5
Owens residence (%)	82.7*	57.4	62.8	68.3
Has had psychotherapy (%)	75.0	82.0	79.1	81.8
First or only child (%)	46.2	50.8	58.1	49.6
Broken family prior to age of 18 (%)	34.6	31.2	25.6	33.9
Judged parents' marriage as happy or very happy (%)	24.5	29.5	25.6	27.8

\* Chi square significant at .01 level.

(peyote, psilocybin, or mescaline); and for self-initiation *vs.* therapist initiation to the use of LSD. Differences were highly significant for usage of marijuana and hashish (chi square = 9.0 and 8.4, respectively). The pattern is

TABLE 2  
DRUG USAGE DATA

Variable	Group			
	I. No accidents or violations (N = 52)	II. Two or more violations (N = 61)	III. One or more accidents (N = 43)	IV. No accidents (N = 121)
Used marijuana	23.1*	52.5	53.5	38.0
Alcohol use $\geq 2 \times$ /day	23.1	32.8	34.5	30.6
Frequent or regular use of				
Sedatives	23.1	18.0	18.6	20.6
Stimulants (amphetamines)	15.4	29.5	21.4	22.3
Tranquilizers	23.1	32.8	32.6	25.6
Nonmedical use of				
LSD	13.5	21.3	25.6	20.6
Hashish	.0*	18.0	11.6	13.2
Opium, morphine, or cocaine	1.9	13.1	11.6	6.6
Heroin	.0	6.6	4.4	4.1
Used peyote or mescaline or psilocybin	9.6	18.0	16.3	14.9
Has injected drugs nonmedically	3.8	11.5	7.0	7.4
Smokes tobacco (currently)	42.3	47.5	46.5	46.3
Introduction to LSD self-initiated	48.0	63.9	48.8	57.8

Note: All figures are percentages.

\* Chi square significant at .01 level.

somewhat similar, although with fewer differences, for the one or more accidents group *vs.* the no accidents group. The one or more accidents group had experimented with or used more of marijuana, alcohol, tranquilizers, non-medical LSD, other strong hallucinogens, opium, morphine, or cocaine; there were virtually no differences in the use of sedatives, stimulants, hashish, and tobacco. Counterintuitively, this group had a lower percentage of subjects who took LSD under experimental (*i.e.*, self-initiated) circumstances rather than at the recommendation of the therapist. For all of the groups except the no accidents or moving violations one, tranquilizers were used more than stimulants, and stimulants were, in turn, used more heavily than sedatives.

### 3. *Personality, Behavioral, and Attitudinal Data*

The no accidents or violations group scored significantly higher on the aphorism scale ( $p < .05$ ) and significantly lower on the relativity of sanity ( $p < .001$ ), sensation-seeking ( $p < .01$ ) [consistent with the results of Conger *et al.* (6) and Goldstein and Mosel (8)], and LSD cultism scale ( $p < .001$ ), than did the two or more violations group (see Table 3). The no accidents or violations group also had a significantly higher rate of attendance at religious services ( $p < .01$ ), in line with the results of Conger and Gaskill (5) and Conger *et al.* (6). Differences approached significance on the two Myers-Briggs Type Indicator subscales for Groups I and II, suggesting that safer drivers prefer a less spontaneous, more ordered, less intuitive mode of life than do those with higher violation rates [congruent with conclusions of Beamish and Malfetti (3), Kole and Henderson (10), McFarland (11), and Tillman and Hobbs (19)]. Persons with no accidents or violations had less belief in paranormal phenomena, were less hypnotically suggestible, were less liberal on social and economic issues and matters of individual freedom (for example, abortion and homosexuality), and were more willing to delay rewards in order to obtain later gratification. The no accidents or violations group also showed more involvement in political activities [similar to Beamish and Malfetti (3)] and organizations of different sorts, but less involvement in meditation, yoga, zen, etc. The differences for the two groups on measures of alienation were negligible.

The only significant difference between the one or more accidents group and the no accidents group was on the sensation-seeking scale ( $p < .01$ ), in the predicted direction.

### 4. *Multiple Regression Analysis*

A multiple regression analysis was done in order to determine what proportion of the driving behavior variance could be accounted for by a combination

TABLE 3  
PERSONALITY, BEHAVIORAL, AND ATTITUDINAL DATA

Variable	Group			
	I. No accidents or violations ( <i>N</i> = 52)	II. Two or more violations ( <i>N</i> = 55)	III. One or more accidents ( <i>N</i> = 39)	IV. No accidents ( <i>N</i> = 112)
<i>Personality</i>				
Aphorism scale	20.5*	22.1	20.5	21.3
Belief in paranormal phenomena	19.2	20.7	19.7	19.9
Myers-Briggs Type Indicator				
Judgment-perception	5.9	7.0	7.2	6.6
Sensing-intuition	8.5	9.3	9.0	8.9
Marlowe-Crowne Social Desirability	7.1	6.7	7.3	6.8
Aas hypnotic suggestibility	54.6	60.3	51.1	57.1
Sensation-seeking	6.2**	7.8	7.9**	6.9
Semantic differential				
Self	84.6	79.9	78.5	82.9
Ideal self	82.8	82.2	79.5	82.9
<i>Attitudes</i>				
Satisfaction with work	13.4	14.2	12.7	13.9
Relativity of sanity	3.4***	4.5	3.6	3.9
Socioeconomic liberalism	26.2	28.3	26.8	27.8
Individual freedom	40.9	45.3	44.6	43.3
Foreign policy liberalism	48.8	52.7	49.8	50.7
Life and death	23.0	25.6	23.8	24.2
Immediate vs. delayed gratification	20.9	22.9	21.2	22.1
Skepticism with legal system	23.7	26.4	25.1	25.1
LSD cultism	4.6***	5.9	4.9	5.0
<i>Alienation</i>				
Srole anomie	3.0	3.2	3.6	3.0
Sociocultural	33.4	33.8	34.8	33.8
Personal	55.8	53.6	54.4	55.3
<i>Participation indices</i>				
Meditation, yoga, zen, etc.	2.4	3.2	2.4	3.0
Organizational involvement	9.2	7.4	6.9	8.5
Political involvement	6.4	5.8	5.8	6.3
Attendance at religious services always or almost always (%)	23.5**	4.9	11.6	14.2

Note: Unless otherwise indicated, all figures are mean values.

\* Significant at .05 level.

\*\* Significant at .01 level.

\*\*\* Significant at .001 level.



of the personality, drug, and attitude variables which discriminated most effectively among the groups. The initial analysis for the four dependent variables (groups) was done with use of all 10 of the independent variables/discriminators. After examination of the beta coefficients, further analyses were done to eliminate the predictors that accounted for the smallest amount of the variance. Only one of the criteria could be very adequately predicted; namely, which subjects had had neither a moving violation nor an accident (the "safe" drivers). Approximately 25% of the variance was accounted for by five predictors: sensation-seeking scale, experimentation with marijuana and hashish, the owning *vs.* the renting of a residence, and sex. The single best predictor was the sensation-seeking score; i.e., how willing a subject was to take chances with unknown experiences, to try new foods, to risk a commission *vs.* a salary, etc. About 10% of the variance was accounted for by this score.

#### D. DISCUSSION

Several conclusions can be drawn from this study. First, it is considerably easier to differentiate "safe" drivers (i.e., those with no moving violations or accidents) from those with two or more violations on the basis of drug, personality, attitudinal, and biographical measures, than it is to differentiate those with accidents from those without. Second, a definite syndrome of personality and behavioral characteristics emerged to describe the person likely to have had several moving violations: he was more liberal in his political and social convictions; more likely to be a male than a female; more likely to rent rather than own his residence; more willing to take risks with his money, time, and life; more inclined to experiment with mood and mind-altering drugs; less likely to be a first or only child; and less likely to be involved in political and organizational activities. A general tendency to tolerate ambiguity and uncertainty in life style, in addition to a dislike for both external and internal control, seems to have been an underlying characteristic. Third, at least limited predictive value was obtained by use of five variables (sex, marijuana and hashish experience, ownership of residence, religious attendance) to predict a "safe" driver.

#### REFERENCES

1. AAS, A. Hypnotizability as a function of non-hypnotic experiences. *J. Abn. Psychol.*, 1963, **66**, 142-150.
2. ADLER, A. *The Practice and Theory of Individual Psychology*. New York: Harcourt, Brace & World, 1927.
3. BEAMISH, J. J., & MALFETTI, J. L. A psychological comparison of violator and non-violator automobile drivers in the 16 to 19 year age range. *Traffic Safety*, 1962, **6**, 12-15.

4. CHAPIN, S. F. *Experimental Designs in Sociological Research*. New York: Harper, 1955.
5. CONGER, J. J., & GASKILL, H. S. Personal and interpersonal factors in motor vehicle accidents. *Amer. J. Psychiat.*, 1957, **113**, 1069-1074.
6. CONGER, J. J., *et al.* Psychological and psychophysiological factors in motor vehicle accidents. *J.A.M.A.*, 1959, **169**, 1581-1587.
7. CROWNE, D. P., & MARLOWE, D. A new scale of social desirability independent of psychopathology. *J. Consult. Psychol.*, 1960, **24**, 349-354.
8. GOLDSTEIN, L. G., & MOSEL, J. N. A factor study of drivers' attitudes with further study on driver aggression. Highway Research Board Bulletin 172, NAS-NRC Publication 532, Washington, D.C., 1958.
9. KENISTON, K. *The Uncommitted*. New York: Harcourt, Brace & World, 1965.
10. KOLE, T., & HENDERSON, H. L. The cartoon reaction scale with special reference to driving behavior. *J. Appl. Psychol.*, 1966, **50**, 311-316.
11. MCFARLAND, R. A. Psychological and psychiatric aspects of highway safety. *J.A.M.A.*, 1957, **163**, 237.
12. MCGLOTHLIN, W. H., & ARNOLD, D. O. LSD Revisited: A ten-year follow-up of medical LSD use. *Arch. Gen. Psychiat.*, 1971, **24**, 35-49.
13. MCGLOTHLIN, W. H., ARNOLD, D. O., & ROWAN, P. K. Marijuana use among adults. *Psychiatry*, 1970, **33**, 433-443.
14. MCGLOTHLIN, W. H., COHEN, S., & MCGLOTHLIN, M. S. Long lasting effects of LSD on normals. *Arch. Gen. Psychiat.*, 1967, **17**, 521-532.
15. MYERS, I. B. *The Myers-Briggs Type Indicator*. Princeton, N.J.: Educ. Test. Serv., 1962.
16. SCHACHTER, S. *The Psychology of Affiliation*, Stanford, Calif.: Stanford Univ. Press, 1959.
17. SMART, R. G., *et al.* Psychoactive drugs and traffic accidents. *J. Safety Res.*, 1969, **1**, 67-73.
18. SROLE, L. Social integration and certain corollaries: An exploratory study. *Amer. Sociolog. Rev.*, 1956, **21**, 709-716.
19. TILLMAN, W. A., & HOBBS, G. E. Accident-prone automobile drivers: Study of psychiatric and social background. *Amer. J. Psychiat.*, 1949, **106**, 321-331.
20. WOODWARD, J. L., & ROPER, E. In Eulau, H., & Eldersveld, S. J. (Eds.), *Political Behavior*. Glencoe, Ill.: Free Press, 1956.
21. ZUCKERMAN, M., *et al.* Development of a sensation-seeking scale. *J. Consult. Psychol.*, 1964, **28**, 477-482.

*Department of Psychology*  
*University of California, Los Angeles*  
*Los Angeles, California 90024*

## CONGRUENCE OF ADOLESCENTS' SELF-CONCEPTS AND PARENTS' PERCEPTIONS OF ADOLESCENTS' SELF-CONCEPTS\*

*Department of Educational Psychology, The University of Georgia; and  
Department of Psychology, Western Carolina University*

---

JOSEPH C. BLEDSOE AND R. GENE WIGGINS

---

### SUMMARY

The purpose of the study was to compare parents' perceptions of their adolescent (ninth-grade) children with perceptions of the adolescents about themselves. Other purposes included comparisons of "misunderstood" and "understood" children, mothers and fathers, boys and girls, and parents' perceptions of how their adolescents perceive themselves.

Ss were 50 boys and 50 girls and their parents from a Southeastern university town of about 50,000 population. The families were of predominantly middle-class background with education and occupation of parents somewhat above-average. Instruments used were an adaptation of Gordon's "How I See Myself" and "How I See My Child" with the addition of one factor (Parent-Teenager Relations) from Offer's "Self-Image Questionnaire."

The data supported the hypotheses that parents perceive their adolescents more favorably than the adolescents perceive themselves; that mothers and fathers agree in their perceptions of their offspring; that parents' perceptions of their adolescents' self-perceptions are more favorable than the adolescents' self-perceptions. Two sex differences, both favorable to girls were obtained on two factors: Teacher-School Relations and Interpersonal Adequacy. Two significant differences on two factors favored "understood" adolescents: Academic Adequacy and Parent-Teenager Relations. In this population, self-perceptions and parent perceptions were generally favorable and relatively congruent.

---

### A. PROBLEM

Combs and Snygg (1) and Sullivan (5) have pointed out that one's conceptions of himself develop as a result of interactions of the self with significant others within his environment. Significant others (for children and adolescents)

---

\* Received in the Editorial Office on November 16, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

include their parents. If parents' perception of their adolescent offspring differ from the adolescent's self-concept, the adolescent becomes aware of these differences and attempts to resolve them in one or more of several ways. If he has a strong need for achievement, he may raise his aspirations. If his need for approval and affection is strong, he may work harder to achieve and please them. If his parents make demands upon him which are perceived as threatening, he may give up and lower his aspirations. Feelings of being misunderstood may thus find expression in a number of ways.

In this study the main purpose was to compare parents' perceptions of their adolescent with perceptions of the adolescent about himself. A second purpose was to analyze perception of being "understood" or "misunderstood" (by parents) in relation to the adolescent's self-concept. Other questions involved comparisons of mothers' and fathers' perceptions, boys' and girls' self-perceptions, and parents' perceptions of how their adolescent offspring perceive themselves.

The following specific research hypotheses were formulated:

1. Parents perceive their adolescent child more favorably than the adolescent perceives himself.
2. Fathers and mothers do not differ in their perceptions of their adolescent offspring.
3. Parents' perceptions of their adolescent's self-perceptions are more favorable than the adolescent's self-perceptions.
4. Boys and girls do not differ in their self-perceptions.
5. Adolescents who state that their parents understand them perceive themselves more favorably than adolescents who state that their parents do not understand them.

## B. METHOD

### 1. *Subjects*

The Ss were 100 ninth graders (50 boys and 50 girls) and their parents. The adolescents were involved in youth programs in 11 churches of a university town in a Southeastern city of approximately 50,000 population. The educational level represented by the parents included six below high school level, 49 high school graduates, 37 with some college, and 108 college graduates. All homes had both parents living together. The socioeconomic grouping of father's occupation, with the 1960 Census Report as a basis for classification, showed 42 in the highest class (#1) followed by 39, 11, seven, and one in each of the succeeding groups, respectively. Mothers included 41 employed and 59 not employed outside the home.



## 2. Instrumentation

The measure of self-concept was an adaptation of Gordon's "How I See Myself" (2) and Offer's "Self-Image Questionnaire" (4). The adapted scale also called "How I See Myself" consists of 50 items based on a five-point Likert scale. The revised scale was adapted for use with parents under the title "How I See My Child," with each corresponding item changed to refer to the child (adolescent) rather than to the "self." Gordon's factor analysis of the scale revealed seven factors (Physical Appearance, Physical Adequacy, Autonomy, Teacher-School Relationships, Academic Adequacy, Interpersonal Adequacy, and Emotions.) Eight items from the Offer scale were included as an eighth factor, a measure of Parent-Teenager Relations. The Gordon scale was devised from categories developed by Jersild (3). Validity based on inferences of observers who observed a sixth-grade child ("Alice") correlated .50 with Alice's actual self-report. Correlations of peers with Alice's self-report averaged .55. Yeatts (6) reported test-retest reliability of the Gordon scale ranging from .78 to .89 after nine days for groups of students in grades 3 through 12 in populations totalling 8979 Ss.

Offer's scale included 11 subscales, only one of which (Parent-Teenager Relations) was used in this study. Internal consistency reliability of this eight-item scale was reported as .86.

Because the eight factors of "How I See Myself" and "How I See My Child" consisted of a varying number of items, each factor score was transformed to a scale from "1" to "5" by dividing the summated score by the appropriate number of items for each factor, thus permitting a direct comparison across the factors. Scores above 3 may be interpreted as "favorable," with scores below 3 as relatively "unfavorable."

The criterion for dichotomizing the Ss as "understood" or "misunderstood" was item #48 of the "How I See Myself" scale. Those who answered "1, 2, or 3" (toward "My parents seldom understand me") were categorized as "misunderstood," and those responding "4" or "5" (toward "My parents understand me very well") were categorized as "understood." Of the Ss, 58 were "understood"; 42 were "misunderstood."

## 3. Procedure

Following contact of the Ss through letters and interviews, the instruments were administered in a central location with adolescents meeting in a separate room from their parents. Data were treated as confidential with code numbers assigned to each S and each parent. Data were analyzed by a series of one-way

analysis of variance, except for the sex and understood-misunderstood variables, which were included in a factorial (two-way) ANOVA. This factorial ANOVA permitted the testing of interaction effects of sex and the understood-misunderstood factor. In addition to the results of the ANOVA, a frequency count of the direction of the differences was made, and the Fisher Exact Probability Test was employed as an additional test of the hypotheses.

### C. RESULTS

Table 1 presents the adjusted means of the various main effect subclasses on the eight dependent variables. That the self-perceptions and parents' perceptions in this population were generally favorable is indicated by the fact that all means were above the "neutral" value of 3.00.

*Hypothesis 1:* In all eight variables, parents' perceptions of their adolescents were more favorable than the adolescents' self-concepts. In two instances, differences were statistically significant: Physical Appearance ( $F_{1, 298} = 21.4$ ;  $p < .001$ ) and Autonomy ( $F_{1, 298} = 6.65$ ;  $p < .05$ ). The Fisher Exact Probability Test indicates that  $p = .004$  for eight of eight differences to favor one group (one-tailed test).

*Hypothesis 2:* That the fathers and mothers perceived their offspring very similarly may be seen by the remarkable congruence of the means. In four variables, the means of the fathers were higher than those of the mothers, and in no instance was the difference statistically significant. The Fisher Test indicates a  $p$  of .50 that means of fathers would be higher than means of mothers.

TABLE 1  
GROUP MEANS ON FACTORS OF "HOW I SEE MYSELF" AND/OR  
"HOW I SEE MY CHILD"

Group	1	2	3	4	5	6	7	8
Parents	3.69	3.97	3.56	3.98	3.84	3.94	3.51	3.68
Adolescents	3.24	3.94	3.36	3.82	3.70	3.83	3.38	3.64
Fathers	3.69	4.02	3.53	3.94	3.85	3.92	3.49	3.69
Mothers	3.68	3.92	3.60	4.03	3.84	3.96	3.53	3.66
Boys	3.24	4.04	3.38	3.61	3.77	3.68	3.46	3.55
Girls	3.24	3.86	3.33	4.01	3.60	3.96	3.27	3.70
Parents' perceptions of adolescents' self-perceptions	3.37	3.89	3.46	3.88	3.70	3.90	3.45	3.56
"Understood"	3.17	3.84	3.40	3.95	3.91	3.87	3.55	3.82
"Misunderstood"	3.31	4.06	3.31	3.68	3.47	3.78	3.18	3.43

<sup>a</sup> The names of the numbered factors are as follows: 1. Physical Appearance, 2. Physical Adequacy, 3. Autonomy, 4. Teacher-School Relationships, 5. Academic Adequacy, 6. Interpersonal Adequacy, 7. Emotions, 8. Parent-Teenager Relations.

*Hypothesis 3:* In all eight variables, parents' perceptions of their adolescents were more favorable than the parents' perception of their children's self-perceptions. In only one instance, however, (Physical Appearance,  $F_{1,998} = 12.2$ ;  $p < .001$ ) was the difference significant. The Fisher Exact Probability Test indicates that eight differences in the same direction would occur by chance ( $p = .004$ , one-tailed test). Although the differences were relatively small, in three instances (Physical Adequacy, Academic Adequacy, and Parent-Teenager Relations) parents' perceptions of their adolescents' self-perceptions were less favorable than the adolescents' self-perceptions.

*Hypothesis 4:* In four variables, boys had more favorable self-concepts than girls (Fisher Exact Test,  $p = .50$ ). In only two instances, however, were the sex differences significant. In Teacher-School Relations ( $F_{1,98} = 6.56$ ;  $p < .05$ ) and in Interpersonal Adequacy ( $F_{1,98} = 8.06$ ;  $p < .01$ ), girls had higher self-concepts than boys. No significant interactions were obtained between the independent variables of sex and understood-misunderstood.

*Hypothesis 5:* In six of eight variables (Fisher Exact Probability = .144, one-tailed test), "understood" adolescents had more favorable self-concepts than "misunderstood" adolescents. In two instances (Academic Adequacy,  $F_{1,98} = 5.23$ ;  $p < .05$ , and Parent-Teenager Relations,  $F_{1,98} = 4.19$ ;  $p < .005$ ) the differences were statistically significant.

#### D. DISCUSSION

The generality of the findings is somewhat limited by the relative homogeneity of the Ss. Even though both parents were present and living together and from predominantly middle-class families, 42 of the 100 ninth-graders (most of whom were 14 or 15 years of age) indicated that their parents seldom understood them. One might suspect that this expression of "misunderstanding" may be a surface manifestation of growing autonomy or that it may be an expression of the popular thing to say, since these adolescents have grown up in a society and at a time (1960's) when rebellion against established mores and "people over 30" are common manifestations. In view of the *Zeitgeist*, it is perhaps remarkable that the perceptions of parents and children are so congruent. The results generally supported Hypotheses 4 and 5. Parents were less critical of their offspring (possibly they saw them as extensions of themselves), but they correctly perceived that their adolescent children were somewhat more critical of themselves. In two instances (Physical Adequacy and Emotions), they thought the children would rate themselves lower than they actually did. These factors may be more visible to the parents, since this period is typically one of rapid growth and change. Two significant sex differences for adolescents

were obtained, both favorable to girls, on two factors: Teacher-School Relations and Interpersonal Adequacy. At this stage of development, girls are usually more mature, physically, socially, and emotionally, and there is apparently less conflict in school situations. Except in these factors, differences were small and sometimes favored boys.

Those adolescents who perceived themselves as "misunderstood" by their parents rated themselves generally lower in the self-concept factors, although only Academic Adequacy and Parent-Teenager Relations were significant. In two instances, "misunderstood" Ss rated themselves higher than "understood" Ss: Physical Appearance and Physical Adequacy. It may be that these areas are simply more important to these "misunderstood" Ss and reflect a different set of values. (They are perhaps perceived as areas of relative ego-strength as compared to the other factors.)

The data revealed that, in general, the teenagers see themselves positively. The parents also view the adolescents in positive ways. The adolescents seem to have accepted the basic roles of society related to academic aspirations and values. These findings are similar to those of Offer (4).

#### REFERENCES

1. COMBS, A. W., & SNYGG, D. *Individual Behavior: A Perceptual Approach to Behavior*. New York: Harper & Row, 1959.
2. GORDON, I. J. *Studying the Child in the School*. New York: Wiley, 1966.
3. JERSILD, A. T. *Child Psychology*. Englewood Cliffs, N. J.: Prentice-Hall, 1960.
4. OFFER, D. *The Psychological World of the Teen-Ager*. New York: Basic Books, 1969.
5. SULLIVAN, H. S. *The Interpersonal Theory of Psychiatry*. New York: Norton, 1953.
6. YEATTS, P. P. An analysis of developmental changes in the self-report of Negro and white children, grades 3-12. Unpublished Doctoral dissertation, University of Florida, Gainesville, 1967.

*Department of Educational Psychology*  
*The University of Georgia*  
*College of Education*  
*Athens, Georgia 30601*



## A STUDY OF THE SELF-ESTEEM AND ALIENATION OF MALE HOMOSEXUALS\*

*State University of New York at Buffalo*

---

JERROLD S. GREENBERG

---

### SUMMARY

The purposes of this study were to determine the differences in self-esteem and alienation between homosexual and heterosexual American males, and the differences in self-esteem and alienation in homosexual males, by selected demographic variables and sexual behavior. Eighty-six American male homosexuals indicated significantly greater alienation than heterosexual males, but similar self-esteem levels. Religion, oral-genital insertion preference, and marital status were significantly related to self-esteem and alienation, whereas anal intercourse insertion preference and degree of homosexuality were not found to be significantly related to self-esteem and alienation.

---

### A. INTRODUCTION

Though much has been written about homosexuality both by homosexuals (4) and heterosexuals (12, 17), few researchers have studied the socially functional American homosexual. Generalizations pertaining to noninstitutionalized homosexuals have been inappropriately drawn from data obtained from homosexuals in clinical (1, 5), military (18), or prison settings. As stated by Clark and Epstein:

One does not go into prisons and mental hospitals or clinics to study normal heterosexuality and then generalize to the population at large; therefore, to do so with homosexuals, as has been done so often in past research dealing with this area, is to study a sample already selectively loaded with psychopathology which may not be truly representative of the whole homosexual population (3, p. 575).

With the development and surfacing of homophile organizations and the gay liberation movement (13), homosexuals in noninstitutionalized settings are available in sufficient numbers for study.

One aspect of the nature of homosexuality that has been studied relates to

---

\* Received in the Editorial Office on November 17, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

the self-concept of the homosexual. However, the evidence relative to self-concept of homosexuals has been contradictory in nature. Whereas some investigators have reported self-acceptance on the part of American homosexuals (2, 9), others have cited depreciated self-esteem (10, 13). Alienation is another variable that has been associated with homosexuality. Researchers have reported greater alienation for homosexuals than for heterosexuals (15, 16). Reports of differences amongst homosexuals relative to self-esteem and alienation, however, have not been located during a search of the literature.

The purposes of this study related to the compilation of data pertaining to self-esteem and alienation on the part of noninstitutionalized male homosexuals. In particular, the following hypotheses were tested:

H<sub>1</sub>: No significant difference exists in alienation for homosexuals and other males.

H<sub>2</sub>: No significant difference exists in self-esteem and alienation by (a) sexual preference, (b) religion, (c) anal insertion preference, (d) fellatio insertion preference, (e) marital status.

H<sub>3</sub>: No significant difference exists in satisfaction with homosexual status for bisexuals and homosexuals.

## B. PROCEDURES

Participation and cooperation of representatives of selected male homophile organizations were solicited and obtained. These representatives requested their membership to participate in this study, and questionnaires were mailed to those members who agreed to cooperate. The outcome of this procedure was a sample of 86 male homosexuals with a mean age of 23.1, enrolled in homophile organizations from such diverse geographic locations as California, Nebraska, New York, and Oregon. The questionnaire to which the participants responded consisted of (a) a self-esteem scale (14) with a range of zero to 6, (b) alienation scale (6) with a range of zero to 96, (c) items pertaining to sexual behavior, and (d) items pertaining to demographic data. Alienation in this study referred to the factors of powerlessness, normlessness, and social isolation.

The subjects were asked to place themselves in a category based upon sexual preference. Although Kinsey has offered a scale for this purpose (11), the fine distinction of the scale was thought to be of possible confusion to the respondents. Therefore, the participants were asked to place themselves in one of the following sexual preference categories: bisexual but predominantly heterosexual, bisexual but predominantly homosexual, or exclusively homosexual.

Although there was concern about the self-report nature of the data collection procedure, the anonymity of the responses and the requests for cooperation by the leadership of the participants' organizations tended to assuage this researcher. However, the findings reported should be considered as limited by the data collection procedure, and interpreted with this limitation in mind.

### C. RESULTS

A mean of 4.60 and standard deviation of 1.78 on the self-esteem scale, when compared with heterosexual males previously measured on the Rosenberg self-esteem scale (8), proved not to be significantly different. However, when compared to Dean's reported mean of 36.64 and standard deviation of 13.5 for heterosexual males, the homosexuals' mean of 45.86 and standard deviation of 16.29 on the alienation scale was found to be different at the .01 level of significance. Though the homosexuals in this study were more alienated than other males, thereby rejecting the first hypothesis, Dean's data were compiled in 1955 (7). Given the more active role of college students in today's society, higher alienation scores than in more subdued times might be expected. It is suggested that future investigations utilize a control group of heterosexual subjects in which alienation measures would be collected and compared to similar measures obtained from an experimental group of homosexual subjects.

When hypothesis 2(a) was tested, as shown in Table 1, there was found to be no significant difference on self-esteem and alienation by sexual preference, although the univariate  $F$  on self-esteem indicated that bisexuals who were predominantly heterosexual had significantly lower self-esteem scores than the other two sexual preference groups. It is hypothesized that bisexuals who were predominantly heterosexual experienced more guilt because of their sexual ambivalence relative to homosexual activities, than other bisexuals or homosexuals, with a resultant depreciation of concept of self.

When the difference in self-esteem and alienation for religion was studied, thereby testing hypothesis 2(b), the  $F$  value of 2.15 proved to be significant, and the hypothesis was rejected. Catholic respondents reported the lowest self-esteem (3.53) and highest alienation (52.13) of the religious groups. The strictness of Catholic dogma relative to homosexuality may have accounted for these findings, since those not aligning with a particular religious group, and thereby not experiencing guilt feelings relative to "sinful" behaviors, reported the second highest self-esteem (4.88) and lowest alienation (43.30) of the religious groups. The role of religion and religious leaders relative to homosexuality and homosexuals needs further study.

TABLE 1  
MULTIVARIATE ANOVA FOR SELF-ESTEEM AND ALIENATION

Source	<i>df</i>	<i>MS</i> <sub>1</sub>	<i>MS</i> <sub>2</sub>	<i>F</i>
By sexual preference and religion <sup>a</sup>				
Sexual preference	2	6.42	258.89	1.69
Religion	3	8.96	358.31	2.15*
Interaction	2**	2.93	309.95	.94
Within group	73	2.12	188.90	
By insertion preferences and marital status				
Anal insertion preference				
Between groups	2	2.07	83.50	.50 (n.s.)
Within groups	77	2.42	180.69	
Fellatio insertion preference				
Between groups	2	8.95	654.92	2.45*
Within groups	82	2.30	180.23	
Marital status				
Between groups	2	.94	305.55	2.39*
Within groups	82	2.47	192.17	

<sup>a</sup> *N* = 81.

\* Significant at the .05 level.

\*\* Four cells were empty.

There were no significant differences in self-esteem and alienation for the interaction of sexual preference and religion.

When hypotheses 2(c) and 2(d) were tested, self-esteem and alienation differences by insertion preferences, only hypothesis 2(d) could be rejected. Those preferring to receive the penis of the partner during oral-genital experiences exhibited the lowest self-esteem (3.73) and highest alienation (55.36) of the fellatio insertion preference groups and all three of the anal insertion preference groups. Those citing no preference relative to fellatio insertion scored highest on self-esteem (4.90) and lowest on alienation (43.83) than the other fellatio insertion preference groups and than all three of the anal insertion preference groups. Since it might be expected that one who prefers to receive the penis of his partner, thereby playing the feminine role in the relationship and identifying to a lesser extent with the masculine role established by the society, would exhibit low self-esteem and high alienation, this finding was not surprising. However, the inability to reject hypothesis 2(c), differences between anal insertion preference groups, remains a perplexity when the difference between fellatio insertion preference groups is considered.

It was expected that those respondents who were married, separated, or divorced at the time of the administration of the testing procedures would



be more alienated than those respondents who never were married. As noted in Table 1, differences in self-esteem and alienation were significant for marital status, and hypothesis 2(e) was rejected. It is significant to note that the never married group had the lowest alienation of all marital status groups. However, the lowest self-esteem of the marital status groups was exhibited by the never married group as well. Since self-esteem and alienation for the total respondents was negatively correlated at the .01 level of significance ( $r = -.63$ ), this finding was unexpected. Perhaps the acceptance of themselves by others to the extent of being married, led to a higher concept of self than for those respondents who never received that degree of acceptance. In evidence of this explanation, the highest self-esteem mean was obtained by those homosexuals who were married at the time of this investigation.

When asked if they would take a pill, if available, which would result in their being exclusively heterosexual, only eight of the 84 homosexuals who responded to this item said they would ingest such a pill. As shown in Table 2, when satisfaction with homosexual status was related to sexual preferences, the  $\chi^2$  value of 2.58 proved not to be significant. Hypothesis 3 could there-

TABLE 2  
RELATIONSHIP BETWEEN SATISFACTION WITH HOMOSEXUAL STATUS  
AND SEXUAL PREFERENCE

Sexual preference	Total response	Satisfaction with status	
		Observed	Expected
Bisexual—predominantly heterosexual	5	4	5
Bisexual—predominantly homosexual	29	25	26
Exclusively homosexual	50	47	50

Note:  $\chi^2 = 2.58$ ,  $df = 2$ ,  $p > .05$ .

fore not be rejected, and it was concluded that there were no differences in satisfaction with homosexual status by sexual preference.

Regarding the responses to the questionnaire, this researcher can only infer from the data and his experience that these responses were indeed reliable and valid. Personal cognizance of homosexual members of homophile organizations has led this writer to conclusions that were suggested by the findings of this investigation. Homosexual males, when presented with the results of this study, have similarly concurred with these findings.

#### D. DISCUSSION

The results of this study seemed to indicate that homosexuals, though feeling alienated from the society in which they lived, were satisfied with

their homosexuality to the extent that they exhibited self-esteem scores that were comparable to heterosexual scores obtained on the same instrument.

The relationship between religion and self-esteem and alienation needs further clarification, as does the relationship between marital experience and self-esteem and alienation. Catholic respondents and those never married exhibited lower self-esteem than others. These relationships are recommended for further study. In addition, why fellatio insertion preference groups differed significantly in self-esteem and alienation, and anal insertion preference groups did not, remains a perplexity.

Lastly, with the development and emergence of homophile organizations, homosexuals who are socially functional are available for study. Generalizations drawn from homosexual samples in clinical, military, or prison settings are no longer necessary; nor have they ever been appropriate. The personality and characterological changes within individual homosexuals from their initial involvement in homophile organizations to their involvement over a period of time should be determined. In this manner the effect of the organization upon the individual can be measured.

#### REFERENCES

1. BIEBER, I., et al. *Homosexuality: A Psychoanalytic Study*. New York: Basic Books, 1962.
2. CHANG, J., & BLOCH, J. A study of identification in male homosexuals. *J. Consult. Psychol.*, 1960, 24, 307-310.
3. CLARK, T., & EPSTEIN, R. Self-concept and expectancy for social reinforcement in noninstitutionalized male homosexuals. *Proc. 77th Ann. Conven. Amer. Psychoanal. Assoc.*, 1969, 4, 575.
4. CORY, D., & LEROY, J. *The Homosexual and His Society: A View from Within*. New York: Citadel Press, 1963.
5. CURRAN, D., & PARR, D. Homosexuality: An analysis of 100 male cases seen in private practice. *Brit. Med. J.*, Apr. 6, 1957, 1, 797-801.
6. DEAN, D. Alienation: Its meaning and measurement. *Amer. Sociolog. Rev.*, 1961, 26, 753-58.
7. DEAN, D. Personal correspondence, Dec. 21, 1971.
8. GREENBERG, J. Masturbation, self-esteem, and other variables. *J. Sex Res.*, in press.
9. HOOKER, E. Forward. In C. Williams & M. Weinberg, *Homosexuals and the Military: A Study of Less than Honorable Discharge*. New York: Harper & Row, 1971.
10. KAPLAN, E. Homosexuality: A search for the ego-ideal. *Arch. Gen. Psychiat.*, 1967, 16, 355-58.
11. KINSEY, A., et al. *Sexual Behavior in the Human Male*. Philadelphia: Saunders, 1948.
12. MARMOR, J. *Sexual Inversion: The Multiple Roots of Homosexuality*. New York: Basic Books, 1965.
13. NEWSWEEK. The militant homosexual. *Newsweek*, Aug. 23, 1971, 45-48.
14. ROSENBERG, M. *Society and the Adolescent Self-Image*. Princeton: Princeton Univ Press, 1965.

15. SCHUR, E. Crimes Without Victims. Englewood Cliffs, N. J.: Prentice-Hall, 1965.
16. SOCARIDES, C. The Overt Homosexual. New York: Grune & Stratton, 1968.
17. WEST, D. Homosexuality. Chicago: Aldine, 1967. P. 138.
18. WILLIAMS, C., & WEINBERG, M. Homosexuals and the Military: A Study of Less than Honorable Discharge. New York: Harper & Row, 1971.

*State University of New York at Buffalo*  
*314 Foster Hall*  
*Buffalo, New York 14214*





## AGE, SEX, AND TITLE OF THERAPIST AS DETERMINANTS OF PATIENTS' PREFERENCES\*<sup>1</sup>

*Southampton College*

---

WILLIAM E. SIMON<sup>2</sup>

---

### SUMMARY

This study investigated age, sex, and title of therapist as determinants of patients' preferences. In general it was found that (a) "psychologists" and "psychiatrists" were preferred to "behavioral consultants," "emotional counselors," and "psychoanalysts" who, in turn, were preferred to "social workers"; (b) male therapists were preferred to female therapists; and (c) 40-year-old therapists were preferred to 55-year-old therapists who, in turn, were preferred to 25-year-old therapists.

---

### A. INTRODUCTION

The basis of patients' preferences for psychotherapists is an important subject area about which clinical psychology has a great deal to learn. All other things being equal, it seems reasonable to predict that the greater the strength of a patient's preference for a therapist, the more effective that therapist will tend to be. There are at least two theoretical reasons for this prediction. First, it would be expected that the greater the strength of a patient's preference for a therapist, the greater the effort he would make, consciously or unconsciously, to communicate to the therapist the nature and source of his distress. Second, it would be expected that the greater the strength of a patient's preference for a therapist, the greater the extent to which he would be influenced by the therapist's communications. That people tend to talk to, and to be influenced by, people whom they like to a greater extent than people whom they dislike is not particularly surprising. The importance of patient preferences is also implicit in a conclusion reached by

---

\* Received in the Editorial Office on November 21, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The author wishes to thank Jo Garber, Carol Ballato, Virginia Leighton, and Anthony Moreno for their assistance in this research. The author is also grateful to a number of students enrolled in the Oceanside adult education program for their help in collecting data.

<sup>2</sup> Requests for reprints should be addressed to the author at the address shown at the end of the article.

Reisman (1) in an interesting review of the literature relevant to the complex question of what makes for a good therapist. Reisman (p. 92) states: "Thus there are no qualities of the therapist that are definitive, and it is preferable to speak, not of a good therapist, but of an appropriate therapist for a particular client."

Given that patient preferences are exceedingly important, what therapist variables serve as the bases for these preferences? Three variables that patients can usually assess immediately, either before or during the first session, are the therapist's age, sex, and title. Unlike so many other important variables that are primarily psychological rather than physical in nature, they do not have to be inferred from behavior. The purpose of the present study was to investigate, within one specific methodological context to be described below, the therapist's age, sex, and title as determinants of patients' preferences.<sup>3</sup>

## B. METHOD

### 1. Subjects

Two samples of Ss were used. In both samples, Ss were asked to give their sex and age. Names, however, were not required. One sample consisted of 102 females who were attending classes at Molloy Catholic College for Women in Rockville Centre, New York. Their mean age was 25.36 years ( $SD = 9.39$ ).

The other sample consisted of 67 Ss recruited by members of an adult education class in clinical psychology taught by the author in Oceanside, New York. There were 33 males and 34 females. Their overall mean age was 36.85 years ( $SD = 13.73$ ). The mean age for the males was 40.27 ( $SD = 14.05$ ), while the mean age for the females was 33.53 ( $SD = 12.75$ ). Oceanside and Rockville Centre are neighboring communities located about 25 miles east of New York City.

### 2. Procedure

Ss were given an instruction sheet which requested them to assume that they had a personal problem that they wanted to discuss with a therapist. Ss were then asked to consider very carefully two hypothetical situations, and to do their best to respond as they believed they would in real life. In

---

<sup>3</sup> During his experience as a psychotherapist, the author has come to believe that these three variables are often important to patients. It is, for example, interesting to note the ideas that some patients have about the differences between psychologists, psychiatrists, and psychoanalysts.

the first hypothetical situation, henceforth referred to as Study I, Ss were asked to assume that there were six therapists who had been equally highly recommended to them and that the only differences among these six therapists that they were aware of were their titles. The instruction sheet went on to inform Ss that therapist A was a behavioral consultant, therapist B was an emotional counselor, therapist C was a psychiatrist, therapist D was a psychoanalyst, therapist E was a psychologist, and therapist F was a social worker. Ss were asked to rank these six therapists from the one whom they believed they would be most likely to consult (to be given a rank of 1) to the one whom they believed they would be least likely to consult (to be given a rank of 6).

In the second hypothetical situation, henceforth referred to as Study II, Ss were asked to assume that there were six therapists who had been equally highly recommended to them and that the only differences among these six therapists that they were aware of were their ages and sexes. The instruction sheet went on to inform Ss that therapist A was a 25-year-old male, therapist B was a 25-year-old female, therapist C was a 40-year-old male, therapist D was a 40-year-old female, therapist E was a 55-year-old male, and therapist F was a 55-year-old female. Ss were asked to rank these six therapists in the same manner as in Study I.

### C. RESULTS AND DISCUSSION

Table 1 shows the mean preference ranks of the six therapists in Studies I and II. In order to investigate age differences in preferences, Ss in each of

TABLE 1  
MEAN PREFERENCE RANKS OF THE SIX THERAPISTS IN STUDIES I AND II

MEAN PREFERENCE RANKS OF THE SIX THERAPISTS IN STUDIES I AND II				
Therapist	Males <i>n</i> = 33	Oceanside sample	Total <i>n</i> = 67	Molloy sample
		Females <i>n</i> = 34		Females <i>n</i> = 102
<i>Study I</i>				
Behavioral consultant	3.88	3.91	3.90	3.71
Emotional counselor	3.61	3.41	3.51	2.85
Psychiatrist	2.36	2.62	2.49	3.35
Psychoanalyst	3.27	3.38	3.33	3.73
Psychologist	2.52	2.65	2.58	2.61
Social worker	5.36	5.09	5.22	4.75
<i>Study II</i>				
25-year-old male	4.39	3.91	4.15	3.68
25-year-old female	5.09	4.56	4.82	3.25
40-year-old male	1.91	2.29	2.10	2.61
40-year-old female	3.36	3.00	3.18	3.13
55-year-old male	2.42	3.32	2.88	3.81
55-year-old female	3.82	3.91	3.87	4.52

the two samples were divided into an Older Group and a Younger Group. The Older Oceanside Group (O-O Group) consisted of all Ss in the Oceanside sample who were at least 40 years of age ( $n = 36$ ), while the Younger Oceanside Group (Y-O Group) consisted of all Ss in the Oceanside sample who were under 40 years of age ( $n = 31$ ). The mean age for the O-O Group was significantly greater than the mean age for the Y-O Group (47.50 *vs.* 24.48,  $t = 12.59$ ,  $df = 65$ ,  $p < .001$ ). The Older Molloy Group (O-M Group) consisted of all Ss in the Molloy sample who were at least 22 years of age ( $n = 36$ ), while the Younger Molloy Group (Y-M Group) consisted of all Ss in the Molloy sample who were under 22 years of age ( $n = 66$ ). The mean age for the O-M Group was significantly greater than the mean age for the Y-M Group (35.94 *vs.* 19.59,  $t = 15.22$ ,  $df = 100$ ,  $p < .001$ ). Table 2 shows the mean preference ranks of the six therapists in Studies I and II for younger and older respondents.

TABLE 2  
MEAN PREFERENCE RANKS OF THE SIX THERAPISTS IN STUDIES I AND II  
FOR YOUNGER AND OLDER RESPONDENTS

Therapist	Oceanside sample		Molloy sample	
	Y-O Group $n = 31$	O-O Group $n = 36$	Y-M Group $n = 66$	O-M Group $n = 36$
<i>Study I</i>				
Behavioral consultant	3.61	4.14	3.48	4.11
Emotional counselor	3.39	3.61	2.53	3.44
Psychiatrist	2.65	2.31	3.82	2.50
Psychoanalyst	3.42	3.25	3.88	3.44
Psychologist	2.68	2.50	2.70	2.44
Social worker	5.26	5.19	4.59	5.06
<i>Study II</i>				
25-year-old male	3.42	4.78	3.35	4.28
25-year-old female	4.16	5.39	2.58	4.50
40-year-old male	2.35	1.89	2.92	2.03
40-year-old female	3.26	3.11	3.11	3.17
55-year-old male	3.52	2.33	4.35	2.83
55-year-old female	4.29	3.50	4.70	4.19

From Tables 1 and 2 it can be seen that psychologists and psychiatrists were generally the most preferred therapists. However, among Ss in the Y-M Group the emotional counselor was the most preferred therapist, while the psychiatrist was a relatively unpopular choice. From Tables 1 and 2 it can also be seen that (a) male therapists were generally preferred to female therapists and that (b) 40-year-old therapists were generally preferred to 55-year-old therapists who, in turn, were preferred to 25-year-old therapists.



The preferences of the Y-M Group were again found to be largely inconsistent with these generally observed preferences.

Additional research is needed to assess the generality of the findings of these two studies. The specific reasons as to why age, sex, and title of therapist have an effect upon patients' preferences is also a question worthy of future investigation. It would, for example, be interesting to replicate Studies I and II, and after Ss have finished ranking the therapists, ask them to describe the reasons for their rankings in as much detail as possible.

#### REFERENCES

1. REISMAN, J. M. *Toward the Integration of Psychotherapy*. New York: Wiley, 1971.

*678 Allwyn Street*

*Baldwin, New York 11510*



## PROCEDURES FOR ESTIMATING MAGNITUDE OF EFFECTS\*

York University, Canada

---

JOHN GAITO AND JOHN FIRTH

---

### SUMMARY

A review is provided of methods that estimate the magnitude of effects within experimental designs. Intraclass correlation type procedures are appropriate with any ANOVA model. Friedman's  $r_m$  and Cohen's power analysis apply to many non-ANOVA type situations, but only to the Fixed Effects ANOVA design. The Gaito utility procedure has two advantages over other intraclass correlation type measures: the coefficients sum to unity (or to 100%), and coefficients for error components are obtained.

---

### A. INTRODUCTION

In tests of the null hypothesis ( $H_0$ ) against the alternative hypothesis ( $H_1$ ), the researcher is concerned with the question of differences. If  $H_0$  is true, the differences are zero; if  $H_1$  is true, the differences are nonzero. However, when  $H_0$  is rejected by conventional statistical procedures, these methods do not indicate *the magnitude of these nonzero differences*. A determination or approximation of these differences would seem to be a useful and important source of information in many studies, including those in which the failure to reject  $H_0$  may be desirable theoretically (20). Magnitude of effect estimates could provide an important function, that of reducing the incidence of trivial "significant" results which are emphasized as important in the literature (3). Thus, in recent years, there has been a reaction against exclusive concern with tests of  $H_0$ , favoring the use of estimates of the statistical sources of variation.

Actually, methods for estimating the magnitude of effects have been available since the early part of the twentieth century, well before the  $F$  and  $t$  tests of  $H_0$  were developed. The purpose of this paper is to review these procedures, with some historical perspective.

---

\* Received in the Editorial Office on November 24, 1972, and published at Provincetown, Massachusetts. Copyright by The Journal Press.

## B. METHODS WITHIN ANOVA PROCEDURES

1. *Historical Influences*

The procedures for estimating magnitude of effects that can remain within the framework of the analysis of variance have been based, directly or indirectly, on two basic concepts: intraclass correlation and expected mean squares [ $E(MS)$ ].

*a. Intraclass correlation.* The essential nature of intraclass correlation consists of a ratio of the variance between experimental units to the total variance (25, 26, 27). This correlation might be determined, for example, for a given characteristic over pairs of twins, the pair being the experimental unit. With increasing homogeneity within units, the within variance would decrease, and the variation between experimental units expressed relative to the total variation (which contains within variance) would increase. With increasing heterogeneity—i.e., with increasing differences within experimental units—the coefficient based on the ratio would decrease. While the original emphasis in the development of this procedure was to derive an index of intraclass homogeneity, the coefficient ( $r_I$ ) can be used as an estimate of the magnitude of the Between Experimental Units source.

*b. Expected mean squares.* The introduction into behavioral research of R. A. Fisher's analysis of variance procedures (13) with the associated  $t$  and  $F$  tests of  $H_0$  eventually stimulated the development of means of estimating the components of variance contained within mean squares ( $MS$ ). Because ANOVA procedures permitted the researcher to vary several variables simultaneously, it was desirable that he should also be able to assess each variable in terms of how much of the total variance it accounted for (a task that the conventional analysis of variance does not perform). The concept of  $E(MS)$ —that is, the expected value of each mean square in a probability sense—formed the basis for several magnitude of effect estimation procedures, including some adaptations of the intraclass correlation (see below). The  $E(MS)$  concept was developed in the 1940's and 1950's by a number of individuals (e.g., 2, 9, 10, 22) and was introduced into psychology during the late 1950's (e.g., 15, 16, 18, 21, 24, 33). When the Mean Square for Within Units is taken as an estimate of  $\sigma_e^2$  (experimental error), the  $E(MS)$  equations for each component can be solved, and the resulting values can be used to form a number of ratios that serve as indices of the size of the associated effect.



## 2. Computational Methods for Deriving Intraclass Correlations

In the early 1900's J. A. Harris reported a simplified method for obtaining a coefficient of intraclass correlation (25, 26, 27). For a simple one-dimensional design, his formulation (27, p. 450) may be modified to be written as formula 1 in Figure 1, where  $k$  is the number of observations in each class,  $S^2$  is the variance of the total set of scores, and  $S_o^2$  is the variance of the class means.

$r_I$  can be estimated from formula 1, or can be obtained from an analysis of variance table, with either the  $MS$  or expected mean squares  $[E(MS)]$  portions of the table. If  $MS$  are used, formula 1 reduces to formula 2 of Figure 1. If  $E(MS)$  are used, the formulation is that of formula 3 in Figure 1. To illustrate, let us assume we have three experimental units and five observations within each unit, with  $MS$  values as given in Table 1. If the  $MS$  formula is used, the value is indicated by formula 2a in Figure 1. Similarly, if the  $E(MS)$  formula is used, the result is indicated by formula 3a in Figure 1.  $\sigma_o^2$  is estimated by  $MS_W$ , or 10;  $\sigma_b^2$  can be then be determined by solving the  $E(MS)$  equation for  $MS_B$ , and the values for  $\sigma_b^2$  and  $\sigma_o^2$  are substituted into formula 3.

TABLE 1  
ANOVA DATA FOR ESTIMATING  $r_I$

Source	SS	df	MS	$E(MS)^a$
Between Units	100	2	50	$\sigma_o^2 + 5\sigma_b^2$
Within Units	120	12	10	$\sigma_o^2$
Total	220	14		

<sup>a</sup> Some individuals use  $\sigma^2$  terms only for  $\sigma_o^2$  and other random effects; others use  $\sigma^2$  terms for all sources because of the greater flexibility that this procedure allows; we will follow this latter method.

As can be seen with formula 3, if there is maximum homogeneity within experimental units,  $\sigma_o^2$  would be 0, and  $r_I$  would equal 1.0. If  $\sigma_b^2$  was 0,  $r_I$  would equal 0. Since, in practice, it is possible to obtain values of  $MS_B$  that are less than those for  $MS_W$ , formula 2 indicates that a negative value for  $r_I$  would ensue; such a value would be set to 0. However,  $r_I$  is normally not computed unless  $H_0$  has been rejected by the  $F$  test, thus precluding the occurrence of negative values.

If there is more than one dimension or factor in an experiment,  $r_I$  for each

$$r_I = \frac{ks_C^2 - S^2}{S^2 (k + 1)} \quad [1]$$

$$r_I = \frac{MS_B - MS_W}{MS_B + (k - 1) MS_W} \quad [2]$$

$$r_I = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_e^2} \quad [3]$$

$$r_I = \frac{50 - 10}{50 + 4(10)} = .44 \quad [2a]$$

$$r_I = \frac{8}{8 + 10} = .44 \quad [3a]$$

FIGURE 1  
FORMULAS 1, 2, 2a, 3, AND 3a

source can be computed as formula 4 (24, p. 60) or formula 5 of Figure 2 (30, p. 199), where  $\sigma_{x_1}^2$  refers to the variance component associated with the source of concern, and  $\sigma_{x_2}^2 \dots \sigma_{x_k}^2$  refers to the variance of the other ( $k - 1$ ) components. The Haggard formulation allows for more than one error component to be represented by  $\sigma_e^2$  (e.g., in calculating  $r_I$  for the fixed effect in a two-factor, Mixed Model design,  $\sigma_e^2$  represents the  $\sigma^2$  component for Within Units and the interaction component); since the Kirk formulation already includes all components in the denominator,  $\sigma_e^2$  refers only to the Within Units component.

Following Harris's papers, intraclass correlation was used in research problems in biology [as noted by Fisher (13), Snedecor (32), and Haggard (24)], but it was not until about 1940 that this measure (or a related one) was introduced into behavioral research. However, since that time, measures similar to intraclass correlation have become common (as the following sections indi-

$$r_I = \frac{\sigma_{x_1}^2}{\sigma_{x_1}^2 + \sigma_e^2} \quad [4]$$

$$r_I = \frac{\sigma_{x_1}^2}{\sigma_{x_1}^2 + \sigma_{x_2}^2 + \dots + \sigma_{x_k}^2 + \sigma_e^2} \quad [5]$$

$$\begin{aligned} & \frac{\sigma_{x_1}^2}{\sigma_{x_1}^2 + \dots + \sigma_{x_n}^2 + \sigma_e^2} + \dots + \frac{\sigma_{x_n}^2}{\sigma_{x_1}^2 + \dots + \sigma_{x_n}^2 + \sigma_e^2} \\ &= \frac{\sigma_{x_1}^2 + \dots + \sigma_{x_n}^2}{\sigma_{x_1}^2 + \dots + \sigma_{x_n}^2 + \sigma_e^2} \end{aligned} \quad [6]$$

$$\frac{p \sigma_s^2}{p \sigma_s^2 + \sigma_e^2} \quad [7]$$

FIGURE 2  
FORMULAS 4, 5, 6, AND 7

cate) for problems analogous to those to which Harris addressed himself; i.e., genetic analyses (7, 31).

### 3. Reliability Estimation

Since the 1940's, psychologists and educators have been using an intraclass correlation type measurement to determine reliability of tests or of experiments. If we consider  $\sigma_{x_1}^2$  of formula 4 as reliable variance and  $\sigma_e^2$  as error variance, then  $r_I$  indicates the proportion of reliable variance present, which is the basic definition of reliability. This measure has been used to estimate the reliability of a test (6, 23, 29) and of an experiment (1, 17). The estimation

of reliability is indicated by formula 6 in Figure 2, where  $\sigma_{x_1}^2 \dots \sigma_{x_n}^2$  represent components of reliable variance, and  $\sigma_e^2$  is error variance. In some cases  $\sigma_e^2$  may represent a number of separate error sources; e.g., subjects, subject interactions, errors of measurement.

In psychological testing, variance due to subjects is considered as reliable variance, usually the only reliable variance of concern. In this case the estimation is  $(MS_S - MS_E)/MS_S$  or, in terms of  $E(MS)$ , as indicated by formula 7 of Figure 2, where  $\sigma_s^2$  and  $\sigma_e^2$  refer to subjects and error variance components, respectively, and  $p$  (which is usually greater than 1) is the number of items used in determining reliability. Notice that the greater the number of such items used, the greater the reliability. Going from the case where the coefficient of  $\sigma_e^2$  is 1 to that in which it is  $p$  is the same as using the Spearman-Brown formula to estimate a test  $p$  times as long as the one for which a reliability has been achieved.

#### 4. Coefficient of Utility (U)

As noted above, Haggard (24) and Kirk (30) have derived formulas for the intraclass correlation suitable for computation from the  $E(MS)$  portion of the ANOVA table. Gaito and Gifford (19) also used a similar measure, based on  $E(MS)$  aspects, in an anthropometric study in order to assess the magnitude of a number of variance components. Estimates of inter- and intrasubject variance, and of inter- and intra-anthropometrist variance, for measurements on each of 11 morphological features were obtained; from these, the percentage of the total variance contributed by each was determined. The results were illustrative of how misleading  $F$  test results may be as a guide in estimating the relative importance of the variance components. For example, although inter- and intra-anthropometrist variance was judged statistically "significant" in a number of features, it was of relatively low magnitude in most of these. For all features, intersubject variance was the most important contributor, varying from about 80% to over 99% of the variance from all sources.

Gaito (16) later introduced this procedure into psychological literature in response to the suggestion by Bolles and Messick (4) that magnitude be determined by a ratio of the sum of squares of the variable of concern to the total sum of squares. Because this latter procedure did not make use of the information contained in the  $E(MS)$  column of the ANOVA table, and thus was not sensitive to the differences in the various ANOVA models (Fixed Effects,

Random Effects, and Mixed Models), it would be difficult to interpret in some cases (16).

This procedure provides a Coefficient of Utility ( $U$ ) which can be demonstrated by the data in Table 1. The  $U$  for  $\sigma_b^2$  is the same as that for  $r_I$  in formula 3 (formula 3b of Figure 3). Unlike traditional  $r_I$  procedures, a coefficient is also obtained for  $\sigma_e^2$  (formula 3c in Figure 3). Both of these coefficients can be expressed as percentages by multiplying by 100.

$$U_b = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_e^2} = \frac{8}{18} = .44 \quad [3b]$$

$$U_e = \frac{\sigma_e^2}{\sigma_b^2 + \sigma_e^2} = \frac{10}{18} = .56 \quad [3c]$$

$$\omega^2 = \frac{SS_B - (b - 1)MS_e}{MS_e + SS_T} \quad [8]$$

FIGURE 3  
FORMULAS 3b, 3c, AND 8

Because  $U$ s can be calculated for all significant components in an experiment, as well as error sources, this procedure possesses the unique feature of producing magnitude estimates that sum to 1 (or 100% if percentages are used). Nonsignificant sources can be ignored, or pooled with the error variance (since the  $MS$  of a nonsignificant source can be considered as another estimate of  $\sigma_e^2$ ).

When more than one dimension is involved,  $U$  is the same as the  $r_I$  defined by Kirk (formula 5), except that only components that have been shown to be present by  $F$  tests are used in either the numerator or denominator.

Endler (11) also used  $E(MS)$  aspects in a fashion similar to that of Gaito in determining the relative contribution of each variance component in three dimensional designs for the Random Effects Model and the Mixed Model. Endler and Hunt (12) applied this procedure in determining the contributions of a number of sources involved in a personality study.



A number of other individuals have suggested the use of similar  $E(MS)$ -based procedures in the formulation of magnitude statements relative to specific problems (e.g., 34).

### 5. $\omega^2$

Another measure that can be used to estimate magnitude aspects, within the Fixed Effects ANOVA model only, is  $\omega^2$ , suggested by Hays (28).  $\omega^2$  can be related to  $E(MS)$ , but it is normally computed from a ratio involving a combination of sums of squares and mean squares. For a main effect ( $B$ ), the formula is 8 in Figure 3, where  $SS_B$  and  $SS_T$  refer to the sums of squares for the  $B$  effect and for Total, respectively;  $MS_e$  is the mean square of the error term, and  $(b - 1)$  is the  $df$  for the  $B$  effect. The values of  $\omega^2$  will be similar to those of  $r_I$  and  $U$ , but will usually be slightly less. For example, with the data of Table 1,  $\omega^2$  is .35 for the Between Units source.

## C. METHODS APPLICABLE OUTSIDE THE ANOVA FRAMEWORK

### 1. Friedman's $r_m$

A recent paper by Friedman (14) provided a more general technique for estimating magnitude effects. The magnitude ( $r_m$ ) is obtained from a table using sample size and the value of the statistical technique (e.g.,  $t$ ,  $z$ ,  $F$ , and  $\chi^2$ ). The  $r_m$  values are determined by means of a point biserial correlation coefficient for  $t$  and  $z$  and by the *eta* and contingency coefficients for  $F$  and  $\chi^2$  tests, respectively. With nonparametric procedures, approximations of  $r_m$  can be obtained.

### 2. Cohen's Power Analysis

The power of a statistical test of  $H_0$  depends on the  $p$  of Type I error ( $\alpha$ ), the size of the sample ( $n$ ), the experimental error, and the actual distance ( $d$ ) between the means of the population distributions from which the samples are drawn (when  $H_1$  is true). Power increases as  $\alpha$  increases,  $n$  increases, error decreases, and distance between means increases. If one is interested in estimating  $d$ , the magnitude of differences, it can be related to the values of  $\alpha$ ,  $n$ , and error with use of power tables. Cohen (8) provided a book indicating this type of analysis for  $t$ ,  $r$ ,  $\chi^2$ , and  $F$  situations.

## D. COMPARISON OF METHODS AND CONCLUSIONS

The Friedman and Cohen methods have a major advantage in that they can be applied to problems for which the ANOVA-based methods are not pertinent. On the other hand, they may yield anomalous results when used in

Random and Mixed Model ANOVAs (5), for they are appropriate only for a Fixed Effects Model.  $\omega^2$  has the most limited application of any of the methods discussed, being applicable only to Fixed Effects ANOVAs.  $U$  and the various  $r_I$  procedures, because they are derived from  $E(MS)$ —which reflects the differences in models—can be used with any ANOVA model. They have the additional advantage of ease of calculation (as does  $\omega^2$ ), since all the necessary information is provided in the ANOVA table and has already been used in performing the  $F$  tests.  $U$  has two advantages over the  $r_I$  methods: the coefficients sum to unity (or to 100%), and error components are included. These aspects facilitate more meaningful comparisons of the various coefficients obtained.

In the beginning of this paper two heuristic arguments in favor of the use of magnitude of effect measures were stated: first, these measures are of interest in their own right and, indeed, may be more intelligible than  $F$  test results on their own; second, a trend toward their frequent use might reduce the number of statistically significant but practically trivial results found in the literature. It is true that these measures may provide the researcher with his single best indicator as to the relative importance of the factors employed in his study. However, in the authors' opinion it would be unwise to calculate these estimates for factors that have *not* been judged to be statistically significant by the  $F$  test, as Grant (20) suggests, since, according to the logic of the ANOVA, failure to reject  $M_0$  indicates that the factors have not been indicated to have either a statistical or a practical significant effect.

It should also be noted that the use of such measures cannot restore the "automaticity of inference" formerly associated with tests of significance (3). Essentially, size of effect measures are descriptive statistics, and the researcher must use some intuitive judgment in interpreting them. There are no *a priori* grounds for labelling a given value of a size of effect estimate as either "large" or "small"; this must be decided within the context of the research. However, in most cases the magnitude of effect measures ought to provide a relatively patent assessment of the relative importance of the variables examined.

#### REFERENCES

1. ALEXANDER, H. W. The estimation of reliability when several trials are available. *Psychometrika*, 1947, **12**, 79-99.
2. ANDERSEN, R. L., & BANCROFT, T. A. Statistical Theory in Research. New York: McGraw-Hill, 1952.
3. BAKAN, D. The test of significance in psychological research. *Psychol. Bull.*, 1966, **66**, 423-437.
4. BOLLES, R., & MESSICK, S. Statistical utility in experimental inference. *Psychol. Rep.*, 1958, **4**, 223-227.

5. BREEN, L., & GAITO, J. Comments on Friedman's  $r_m$  procedure. *Psychol. Bull.*, 1970, **73**, 309-310.
6. BURT, C. Test reliability by analysis of variance. *Brit. J. Stat. Psychol.*, 1955, **8**, 103-118.
7. ———. Inheritance of general intelligence. *Amer. Psychol.*, 1972, **27**, 175-190.
8. COHEN, J. *Statistical Power Analysis for the Behavioral Sciences*. New York: Academic Press, 1969.
9. CRUMP, S. L. The present status of variance component analysis. *Biometrics Bull.*, 1946, **2**, 7-11.
10. DANIELS, H. E. The estimation of components of variance. *J. Royal Stat. Soc.*, 1939, Suppl. 6, 186-197.
11. ENDLER, N. S. Estimating variance components from mean squares for random and mixed effects analysis of variance models. *Percept. & Motor Skills*, 1966, **22**, 559-570.
12. ENDLER, N. S., & HUNT, J. McV. Sources of behavioral variance as measured by the S-R inventory of anxiousness. *Psychol. Bull.*, 1966, **65**, 336-346.
13. FISHER, R. A. *Statistical Methods for Research Workers*. Edinburgh, Scotland: Oliver & Boyd, 1925.
14. FRIEDMAN, H. Magnitude of experimental effect and a table for its rapid estimation. *Psychol. Bull.*, 1968, **70**, 245-251.
15. GAITO, J. The single latin square design in psychological research. *Psychometrika*, 1958, **23**, 369-378.
16. ———. The Bolles-Messick coefficient of utility. *Psychol. Rep.*, 1958, **4**, 595-598.
17. ———. Nonparametric methods in psychological research. *Psychol. Rep.*, 1959, **5**, 115-125.
18. ———. Expected mean squares in analysis of variance procedures. *Psychol. Rep.*, 1960, **7**, 3-10.
19. GAITO, J., & GIFFORD, E. C. Components of variance in anthropometry. *Hum. Biol.*, 1958, **30**, 120-127.
20. GRANT, D. A. Testing the null hypothesis and the strategy and tactics of investigating theoretical models. *Psychol. Rev.*, 1962, **69**, 54-61.
21. GREEN, B. F., & TUKEY, J. W. Complex analyses of variance: General problems. *Psychometrika*, 1960, **25**, 127-152.
22. GREENWOOD, J. A. Analysis of variance and components of variance—Factorial experiments. Unpublished paper, Bureau of Aeronautics, Washington, D. C., 1956.
23. GUILFORD, J. P. *Psychometric Methods*. New York: McGraw-Hill, 1954.
24. HAGGARD, E. A. *Intraclass Correlation and the Analysis of Variance*. New York: Dryden Press, 1958.
25. HARRIS, J. A. On the formation of correlation and contingency tables when the number of combinations is large. *Amer. Natural.*, 1911, **45**, 566-571.
26. ———. The formation of condensed correlation tables when the number of combinations is large. *Amer. Natural.*, 1912, **46**, 477-486.
27. ———. On the calculation of intraclass and interclass coefficients of correlation from class moments when the number of possible combinations is large. *Biometrika*, 1913, **9**, 446-472.
28. HAYS, W. L. *Statistics for Psychologists*. New York: Holt, Rinehart & Winston, 1963.
29. HOYT, C. Test reliability obtained by analysis of variance. *Psychometrika*, 1941, **6**, 153-160.
30. KIRK, R. E. *Experimental Design Procedures for the Behavioral Sciences*. Belmont, Calif.: Wadsworth, 1968.
31. ROBERTS, R. C. Some concepts and methods in quantitative genetics. In J. Hirsch (Ed.), *Behavior-Genetic Analysis*. New York: McGraw-Hill, 1967.

32. SNEDECOR, G. W. *Statistical Methods*. Ames, Iowa: Iowa State Coll. Press, 1946.
33. STANLEY, J. Fixed, random, and mixed models in the analysis of variance as special cases of finite model III. *Psychol. Rep.*, 1956, **3**, 369.
34. VAUGHAN, G. M., & CORBALLIS, M. C. Beyond tests of significance: Estimating strength of effects in selected anova designs. *Psychol. Bull.*, 1969, **72**, 204-213.

*Department of Psychology*  
*York University*  
*4700 Keele Street*  
*Downsview 463, Ontario*  
*Canada*





## THE DYSLEXIC CHILD—TWO YEARS LATER\* <sup>1, 2</sup>

*Navy Medical Neuropsychiatric Research Unit and San Diego  
City Schools Guidance Center*

EDNA J. HUNTER<sup>3</sup> AND HADLEY M. LEWIS

### SUMMARY

A follow-up was made of 18 male dyslexic children (RDs) and a group of matched controls (Cs), who had all been judged free from emotional and school adjustment problems two years earlier, (a) to see if RDs' reading deficit had been overcome, (b) to measure "emotional overlay," and (c) to find predictors of success in learning to read. Results showed that at the time of the follow-up (a) none of the RDs had overcome his reading disability despite remedial efforts; (b) RDs had significantly lower achievement levels and significantly more adjustment problems than Cs; (c) low scores for WISC Full Scale IQ, Vocabulary subtest, and Coding subtest were associated with large reading deficits two years later; and (d) high WISC Arithmetic and Coding subtest scores were associated with large gains in reading skill. Physiological measures did not predict size of deficit or gain in reading skill over time.

### A. INTRODUCTION

The importance of early diagnosis of the potential dyslexic child to prevent permanent reading disability and to eliminate the emotional overlay which accumulates with repeated failure within the school environment has often been stressed. The present study was a two-year follow-up of 18 male non-readers, aged 9 years 6 months to 13 years 2 months, and a control group of average or above-average readers (matched by group) who had taken part in an earlier study (1, 2, 3). At the time of the initial study, parental and

\* Received in the Editorial Office on November 27, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was accomplished at the U. S. Naval Hospital, San Diego, and was supported by Department of the Navy, Bureau of Medicine and Surgery, under Task No. MR006.02. The opinions and assertions contained herein are the private ones of the writers and are not to be construed as official or as reflecting the views of the Navy Department.

<sup>2</sup> The writers express their appreciation to Dr. Laverne C. Johnson for his helpful comments and to Ann Clay for typing the manuscript.

<sup>3</sup> Direct requests for reprints to the first author at the address shown at the end of this article.

teacher reports and ratings showed that none of the children had appreciable emotional or behavioral adjustment problems within the home or school situation.

The purpose of the two-year follow-up was threefold: (a) to see if the large reading deficit of the nonreader had increased, decreased, or remained relatively stable over time; (b) to measure the "emotional overlay," if any, which had developed during the two-year interim; and (c) to discover which psychological, physiological, or developmental factors evident two years earlier might have predicted which nonreaders would eventually overcome the reading disability and which ones would not.

## B. METHOD

### 1. Subjects

Subjects were 18 of the 20 male reading disabled (RDs) children (e.g., children who demonstrated a one year or greater deficit in reading ability) who had taken part in the initial study and the 20 control (Cs) subjects of the previous study. At the time data were collected for the follow-up study, mean age for RDs was 11 years 7 months; for Cs, 11 years 8 months. Ages ranged from 9 years 6 months to 13 years 5 months ( $N = 38$ ). Mean grade level for RDs was 6.1; for Cs, 6.5.

### 2. Procedure

Data obtained for each of the 38 children included scores on a recent test of reading skill (Wide Range Reading Test of the Wide Range Achievement Test) and teacher ratings of school behavior, academic achievement, and emotional adjustment. The Wechsler Intelligence Scale for Children (WISC) was not readministered for the follow-up; WISC scores from the earlier study were used for all analyses.

## C. RESULTS

### 1. Intellectual Levels

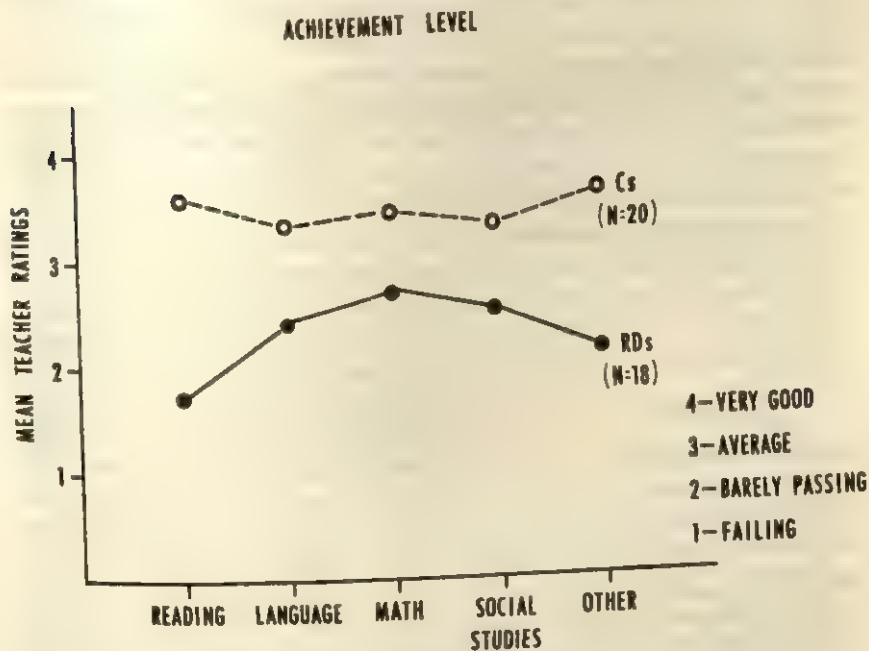
The RD and C groups were not significantly different for WISC Full Scale  $IQ$  (Full Scale  $IQ = 115.3 \pm 10.2$ ,  $N = 38$ ) or WISC Performance  $IQ$  (Performance  $IQ = 111.4 \pm 10.5$ ,  $N = 38$ ). The two groups differed significantly, however, on WISC Verbal  $IQ$  (two-tailed  $t$  test for independent means,  $p < .001$ ). Mean Verbal  $IQ$  for RDs was  $110.0 \pm 10.4$ ; for Cs,  $122.5 \pm 10.3$ .

## 2. Achievement Levels

On the basis of teacher ratings, there were statistically significant differences between the groups at the time of the follow-up with respect to achievement in all school subjects, as shown in Figure 1. These differences reached the .01 level of significance for mathematics and social studies, and the .001 level for reading and language achievement (two-tailed  $t$  tests for uncorrelated means).

## 3. Reading Skills—Two Years Later

Of the 18 RD children in the follow-up, all but one child had received remedial assistance in language skills of some kind during the two-year interim; e.g., special classes in remedial reading, classes for the educationally handicapped, individual instruction, etc. In spite of remedial efforts, *not one of the 18 children had overcome his reading disability*. All RDs continued to demonstrate reading deficits equal to one year or more below expected age-



**FIGURE 1**  
DIFFERENCES IN MEAN ACHIEVEMENT LEVELS IN SCHOOL SUBJECTS FOR NONREADERS (RDs) AND MATCHED CONTROLS (Cs) BASED UPON TEACHER RATINGS

grade reading level. Two years earlier, mean reading deficit for RDs was 2.5 years; data from the follow-up showed a mean deficit of 2.9 years, with deficits ranging from 1.0 years to as high as 5.5 years.

Data from the original study had shown statistically significant differences between readers and nonreaders for scores on the Bender Visual Gestalt Test; the WISC Verbal *IQ*; the WISC Information, Arithmetic, Similarities, Vocabulary, Digit Span, Picture Completion, and Coding subtests; tested laterality; familial incidence of dyslexia; birthdate in relation to school year; motor reaction time; skin conductance level; number electrodermal offset responses; and heart rate variability. Full reports of these findings appear elsewhere (1, 2, 3).

For nonreaders, correlation coefficients between absolute gains in reading skills made in the two-year period and measures taken two years previously showed that the only measures with predictive validity were the WISC Arithmetic subtest score and the WISC Coding subtest score (see Table 1). Subjects with higher scores showed the greatest reading gains. Those RDs

TABLE 1  
CORRELATIONS BETWEEN PSYCHOPHYSIOLOGICAL MEASURES FOR NONREADERS TAKEN TWO YEARS PREVIOUSLY AND GAIN IN READING SKILL AND MAGNITUDE OF READING DEFICIT TWO YEARS LATER

Measure	Reading gain	Reading deficit
<i>Test measures</i>		
Bender Gestalt Visual Motor Test	.06	.15
WISC Full Scale <i>IQ</i>	-.19	-.45*
WISC subtests		
Information	-.16	-.31
Vocabulary	.35	-.58**
Arithmetic	.60**	-.36
Similarities	-.30	.02
Digit Span	-.04	-.07
Picture Completion	.13	-.08
Coding	.54**	-.65**
<i>Physiological measures</i>		
Motor reaction time		
Electrodermal response	-.08	.35
(number offset responses)		
Skin conductance level	-.02	.15
(decrease over time)		
Heart rate variability	-.22	.32
(standard deviation of the beat-by-beat heart rate)	.28	.26

Note:  $N = 18$ .

\*  $p < .05$ .

\*\*  $p < .01$ .

with the highest WISC Full Scale  $IQ$  scores did not necessarily make the largest gains in reading skill over time. However, subjects with lower Full Scale WISC  $IQ$ s at fourth grade level had larger reading deficits present at sixth grade level ( $r = -.45$ ,  $p < .05$ ,  $N = 18$ ). Larger reading deficits were also associated with lower scores on the WISC Vocabulary subtest ( $r = -.58$ ,  $p < .01$ ,  $N = 18$ ) and on the WISC Coding subtest ( $r = -.65$ ,  $p < .01$ ,  $N = 18$ ). None of the physiological measures taken at the time of the earlier study correlated significantly with magnitude of the reading deficit two years later.

At the time of the initial study the younger half of the RD group did not differ from the older half with respect to magnitude of the reading deficit. Mean deficit for younger RDs was 2.7 years, as compared with a deficit of 2.3 years for the older RDs. However, two years later, mean reading deficit for the younger half was 3.8 years; for the older RDs, 2.0 years. This difference between the younger and older nonreaders was statistically significant at the .001 level (two-tailed  $t$  test for uncorrelated means).

As expected, Cs were still significantly better readers than RDs two years after the initial study (two-tailed  $t$  test for independent means,  $p < .001$ ). At the time of the follow-up, Cs showed a mean reading acceleration above expected age-grade level of 1.7 years. Acceleration shown two years earlier was 1.9 years. Reading levels for the majority of Cs were approximately equal to those demonstrated two years previously, although one control subject (Full Scale WISC  $IQ = 112$ ) would now be categorized as "reading disabled"; that is, he was at least one year below expected age-grade level. Three control subjects were even more highly accelerated than two years previously; two tested more than five years above expected age-grade level in reading skill. (Mean Full Scale WISC  $IQ$  for those two Ss was 130.0, which was considerably higher than the mean for the group as a whole.)

#### 4. *Did an Emotional Overlay Develop over Time?*

For the original study, only those children were included who showed no apparent emotional problems and were not behavioral problems within the school situation on the basis of parental and teacher reports and personal interviews. Data from the first study showed only two "adjustment" measures on which RDs differed significantly from the control group. The RDs were rated as less self-confident than Cs by their mothers and were rated by the experimenter as displaying more hyperactivity during the laboratory session.

Two years later, however, there were a large number of differences between the two groups with respect to emotional and school adjustment problems



## CLASSROOM BEHAVIOR

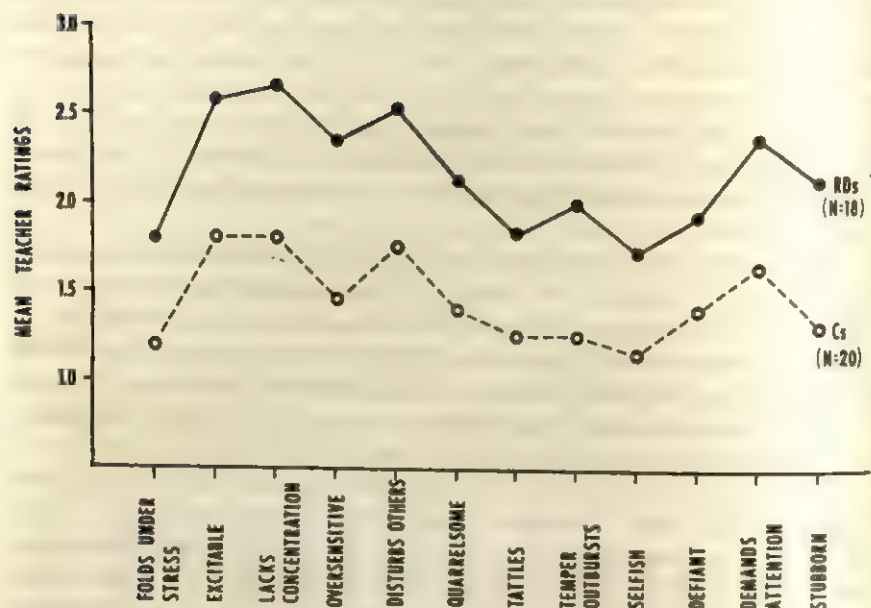


FIGURE 2

THE DEGREE TO WHICH VARIOUS CLASSROOM BEHAVIORS WERE DISPLAYED BY NONREADERS (RDs) AND A MATCHED CONTROL GROUP (Cs)  
BASED UPON TEACHER RATINGS

Ratings were as follows: 4 = Excessive; 3 = Often; 2 = Occasionally; and 1 = Never. All between-group differences were statistically significant at the .05 level or better.

(see Figure 2). On the basis of scores from a teacher rating scale, there were highly significant differences between RDs and Cs with respect to the ability to concentrate, oversensitivity, temper outbursts, and stubbornness (two-tailed  $t$  test for independent means,  $p < .01$ ). The RD child was also judged significantly more excitable, selfish, quarrelsome, defiant, and attention-demanding. He "tattled" more, tended to be a greater disturbance to other children, and was more apt to "fall apart" under stress (two-tailed  $t$  test for independent means,  $p < .05$ ).

## D. DISCUSSION

The present study was a two-year follow-up of 18 reading disabled male children, aged 9 years 6 months to 13 years 2 months, and a matched control group. Scores on the Wide Range Reading Test, school achievement scores

in related subjects, and behavioral ratings by teachers showed that during the two-year interval, in no instance was the nonreader's deficit overcome, and an "emotional overlay" was apparent for the RDs which was not present for the controls. The teacher rating scale used for the follow-up was not identical with that used for the initial study. Therefore, it is possible that these adjustment and emotional differences between the groups were present to some extent two years earlier.

Correlation coefficients between magnitude of the reading deficit at the time of the follow-up study and measures taken two years earlier indicate that low scores on the WISC Full Scale IQ, the WISC Vocabulary subtest, and the WISC Coding subtest were significantly related to large reading deficits two years later. The WISC Arithmetic subtest score and Coding subtest score were significantly and positively related to increases in reading ability over time (WISC Arithmetic,  $r = +.60$ ; WISC Coding,  $r = +.54$ ;  $p < .01$ ). None of the physiological measures (motor reaction time, skin conductance level, electrodermal offset response, heart rate variability) was found to be related significantly to future reading deficit or to increase in reading skill over the two-year period.

The mean reading deficit of the older half of the RD group was approximately equal in magnitude to the mean for that half found at the time of the initial study. The younger half of the RD group showed a significantly larger mean deficit than was evident previously. In the selection of subjects for the initial study, children were considered as reading disabled if demonstrated reading skills fell at least one full year or more below expected age-grade level. It may be that the earlier a reading deficit is recognized, the more serious the deficit. In other words, a one-year deficit at third grade level is a greater deficit than a one-year deficit at fifth grade level.

In a recent study, Maginnis (4) found that those children with the largest reading deficits are not necessarily the ones who will profit most from remediation. Data from this follow-up study showed a negligible correlation ( $r = +.05$ ,  $N = 18$ ) between the degree of reading disability present two years earlier and the size of gain in reading skills after two years. Retrospectively, the correlations between reading deficit at time of follow-up and the gain in skill over the preceding two years was  $-.60$  ( $p < .01$ ,  $N = 18$ ). Not one of the dyslexic children had overcome his reading disability, and nearly half the group ( $N = 8$ ) made only minimal progress—in spite of special remedial procedures. In many cases, the child who was reading at second grade level when he was in the fourth grade was still reading at approximately the second grade level when he was in the sixth grade. It is not

surprising that serious emotional and school adjustment problems appeared to develop over the two-year period.

Silberberg and Silberberg (5) have pointed to seven recent studies which indicate that children subjected to remedial reading show only small short-term positive effects. Those authors emphasized that reading is but a "tool" to educate, and when it becomes evident that a child is unable to learn to read, perhaps another educative tool should be considered. Results of the present study show that the "emotional cost" of not learning to read is exceedingly high for the dyslexic child, and the "return" on remedial investment exceedingly low. We need another tool!

#### REFERENCES

1. HUNTER, E. J. Autonomic responses to aircraft noise in dyslexic children. *Psychol. in Sch.*, 1971, 8, 362-367.
2. HUNTER, E. J., & JOHNSON, L. C. Developmental and psychological differences between readers and nonreaders. *J. Learn. Disabil.*, 1971, 4, 572-577.
3. HUNTER, E. J., JOHNSON, L. C., & KEEFE, F. B. Electrodermal and cardiovascular responses in nonreaders. *J. Learn. Disabil.*, 1972, 5, 187-197.
4. MAGINNIS, G. H. Reading disability and remedial gain. *J. Learn. Disabil.*, 1971, 4, 322-324.
5. SILBERBERG, N. E., & SILBERBERG, M. C. The bookless curriculum: An educational alternative. *J. Learn. Disabil.*, 1969, 2, 302-307.

*Navy Medical Neuropsychiatric Research Unit  
San Diego, California 92152*

# ABBREVIATIONS OF WORDS USED IN THE TITLES OF JOURNALS

(One-word titles are never abbreviated. The word "the" is not used, nor its equivalent in any other language. The word "of" or its equivalent in other languages is used only to discriminate what would otherwise be identical titles in different languages. The word "and" is always used, but indicated by "&" in the Roman alphabet. Only English words are indicated here, but the corresponding words in other languages should receive a corresponding abbreviation. All abbreviations and all one-word titles should be in italics.)

Abnormal .....	<i>Abn.</i>	Japanese .....	<i>Jap.</i>
Abstracts .....	<i>Abst.</i>	Journal .....	<i>J.</i>
American .....	<i>Amer.</i>	Mathematical .....	<i>Math.</i>
Anatomy .....	<i>Anat.</i>	Measurement .....	<i>Meas.</i>
Animal .....	<i>Anim.</i>	Medical .....	<i>Med.</i>
Applied .....	<i>Appl.</i>	Mental .....	<i>Ment.</i>
Archives .....	<i>Arch.</i>	Monographs .....	<i>Monog.</i>
Association .....	<i>Assoc.</i>	Neurology .....	<i>Neurol.</i>
Attitude .....	<i>Attit.</i>	Opinion .....	<i>Opin.</i>
Australian .....	<i>Aust.</i>	Orthopsychiatry .....	<i>Orthopsychiat.</i>
Behavior .....	<i>Behav.</i>	Personality .....	<i>Personal.</i>
British .....	<i>Brit.</i>	Personnel .....	<i>Person.</i>
Bulletin .....	<i>Bull.</i>	Philosophy .....	<i>Philos.</i>
Bureau .....	<i>Bur.</i>	Physics .....	<i>Phys.</i>
Canadian .....	<i>Can.</i>	Physiology .....	<i>Physiol.</i>
Character .....	<i>Charac.</i>	Proceedings .....	<i>Proc.</i>
Children .....	<i>Child.</i>	Psychiatry .....	<i>Psychiat.</i>
Chinese .....	<i>Chin.</i>	Psychoanalysis .....	<i>Psychoanal.</i>
Clinical .....	<i>Clin.</i>	Psychology .....	<i>Psychol.</i>
College .....	<i>Coll.</i>	Psychosomatic .....	<i>Psychosomat.</i>
Comparative .....	<i>Comp.</i>	Quarterly .....	<i>Quart.</i>
Consulting .....	<i>Consult.</i>	Religious .....	<i>Relig.</i>
Contributions .....	<i>Contrib.</i>	Research .....	<i>Res.</i>
Development .....	<i>Devel.</i>	Review .....	<i>Rev.</i>
Educational .....	<i>Educ.</i>	School .....	<i>Sch.</i>
Experimental .....	<i>Exper.</i>	Science .....	<i>Sci.</i>
General .....	<i>Gen.</i>	Social .....	<i>Soc.</i>
Genetic .....	<i>Genet.</i>	Statistics .....	<i>Stat.</i>
Indian .....	<i>Ind.</i>	Studies .....	<i>Stud.</i>
Industrial .....	<i>Indus.</i>	Teacher .....	<i>Teach.</i>
International .....	<i>Internat.</i>	University .....	<i>Univ.</i>
Italian .....	<i>Ital.</i>		

# Preparation of Manuscripts for The Journal Press

21 JUN 1973

## GENERAL INSTRUCTIONS

1. The proper sequence for the parts of your submitted manuscript is as follows: (a) text (b) references, (c) footnotes, (d) tables, (e) figures, and (f) figure legends. However, monographs start with a table of contents and may have an acknowledgment page before the text and an appendix immediately after the text.
2. Use heavy typewriter paper,  $8\frac{1}{2} \times 11$  inches, double-space all lines, and leave margin for editorial work. Do not use onionskin or odd sizes of paper.
3. Submit original typewritten copy and one carbon. Retain second carbon for proofing.
4. Retype any page on which written corrections have been made.
5. Do not begin a sentence with a numeral.
6. An adequate summary serving as the conclusion of the text is good form.
7. Each quotation should indicate the page number of the original source. The original publisher must give permission for lengthy quotations and use of tables or figures.
8. Do not fold your manuscript.
9. Enclose your covering letter with your manuscript. If you are unknown to the Editors, kindly state your credentials.

## FORMAT AND SPECIFIC INSTRUCTIONS

### A. TEXT DIVISIONS

#### I. THE TITLES OF JOURNAL ARTICLES AND THE MAJOR SUBDIVISIONS OF MONOGRAPHS ARE PRINTED IN TEN-POINT CAPS CENTERED ON THE PAGE

##### A. THE NEXT SUBDIVISION TITLE IS PRINTED IN CAPS AND SMALL CAPS CENTERED ON THE PAGE

1. Then *Italics*, with Principal Words, Upper and Lower Case, Centered on the Page
  - a. Then *italics*, upper and lower case, 1-em run-in side head.
    - (1). Then *italics*, upper and lower case, 2-em run-in side head.
      - (a). Then *italics*, upper and lower case, 3-em run-in side head.
- [Further subdivision should be merged into the text without marginal indentation, and should be numbered with small letters.]

### B. REFERENCES

References should be arranged in alphabetical order by author, numbered and referred to in the text by number (2)

The proper form of a book reference is as follows:

2. Doe, J. *The Preparation of Manuscripts*. New York: Holt, 1963. Pp. 400-418.

The proper form of a journal reference is as follows:

2. Doe, J. The preparation of manuscripts. *J. Gen. Psychol.*, 1963, 68, 450-462.

If in the text it is desirable to refer to a page, thus (2, p. 45).

### C. FOOTNOTES

Use as few as possible and number consecutively in the text thus,<sup>1</sup>

<sup>1</sup> Footnote (on the separate footnote page).

### D. TABLES

Tables should always be typed on separate sheets of paper and should be Arabic numbered (Table 2). Construction should be such that vertical lines are unnecessary. Always indicate in the text the approximate insertion place for a table.

### E. FIGURES

Figures should be submitted as glossy prints of the approximate size for final reproduction (ordinarily 4-4½ inches in width). Lettering and lines should be sharp and clean. Figures should be Arabic numbered (Figure 2) in the text, with the approximate insertion point indicated. Each figure requires a legend, but all legends should be submitted double-spaced on a separate Figure Legends page.



March 2, 1973  
19/6/73  
\$39.00 per annum  
\$18.00 per volume  
Single Numbers, \$9.00

BIMONTHLY  
Three volumes per year  
Immediate publication

March, 1973  
Volume 83, Second Half

Founded in 1935 by Carl Murchison

# The Journal of PSYCHOLOGY

The General Field of Psychology

POWELL MURCHISON  
*Managing Editor*

HELEN M. DAVIS  
*Associate Editor*

## Editorial Board

JOHN A. ANTOINETTI  
Life Insurance Agency  
Management Association

S. HOWARD BARTLEY  
Michigan State University

ARTHUR G. BILLS  
University of Cincinnati

E. A. BOTT  
University of Toronto

LEONARD CARMICHAEL  
National Geographic Society

IRVIN L. CHILD  
Yale University

HERBERT S. CONRAD  
U. S. Office of Education

JOHN F. DASHIELL  
Florida State University

LEONARD W. DOOB  
Yale University

If this space should be unstamped, this is the regular library edition.  
But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not  
be resold for a period of five years from date of publication.

FRANK A. GILDARD  
Princeton University

ERNEST R. HILGARD  
Stanford University

JOHN E. HORROCKS  
Ohio State University

DAVID HOTHERBALL  
Ohio State University

HEINRICH KLEVER  
University of Chicago

NORMAN L. MUNN  
University of Adelaide,  
Australia

WARREN H. TRICKNER  
New Mexico State  
University

LEE EDWARD TRAVIS  
Fuller Theological  
Seminary

JOSEPH ZUBIN  
Psychiatric Institute,  
Columbia University

Copyright © 1973 by The Journal Press  
Provincetown, Massachusetts, U. S. A.

All rights reserved

Second-class postage paid at Provincetown, Mass.  
and additional mailing offices

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

## Back Issues

The Journal Press maintains a complete inventory of back issues for all five journals by reprinting many of them each year. Early volumes frequently consist of multiple parts, and the cost of maintaining these volumes is naturally much higher than that for two-part volumes. So our old policy of a flat price per volume cannot be continued. The following rates are in effect: single numbers, \$9.00; two-part volumes, \$18.00; three-part volumes, \$27.00; four-part and over-four-part volumes, \$36.00. The only exceptions (special issues) are starred below.

	VOLUMES	DATES	PRICE PER VOLUME
<i>Journal of Genetic Psychology</i> (Pedagogical Seminary)			
	1-4	1891-April 1897	\$27.00
	5-37	July 1897-1930	36.00
	38*	1930	18.00
	39	1931	36.00
	40 and later	1932-	18.00
<i>Genetic Psychology Monographs</i>			
	1-22	1926-1940	36.00
	23 and later	1941-	18.00
<i>Journal of General Psychology</i>			
	1-3	1928-1930	36.00
	4*	1930	18.00
	5	1931	36.00
	6 and later	1932-	18.00
<i>Journal of Social Psychology</i>			
	1-10	1930-1939	36.00
	11 and later	1940-	18.00
<i>Journal of Psychology</i>			
	1 and later	1936-	18.00

A mailing charge of \$0.30 per volume is made for foreign deliveries only. Domestic mailings are prepaid. All remittances should be by check or money order payable in U.S. funds on a United States bank. Please do not send UNESCO coupons. Advance payment is required from all individuals and from foreign purchasers. Microfilm versions and special bindings are not available.

THE JOURNAL PRESS

2 Commercial Street

Provincetown, Massachusetts 02657

U. S. A.

*Notice: All copyrighted material reprinted from The Journal Press publications should be given an appropriate acknowledgment. No permission is necessary for brief quotations (less than 300-word total or less than 10 percent of length, whichever is least, for any one article). Permission is required for reprinting tables, figures, entire articles, or substantial portions thereof; and such permission will be contingent upon agreement by the original author or authors. Estate permission may be required. Since no fee is charged for a copyright license, the courtesy of a stamped, self-addressed envelope will be appreciated.*

---

# The Journal of Psychology

Founded in 1935 by Carl Murchison

The Journal of Psychology was founded at the request of many psychologists throughout the world for the purpose of establishing a medium in which publication shall be immediate and in which other publication experiments may be carried out.

(1). The subscription price is \$39.00 per annum, \$18.00 per volume, \$9.00 per issue (three volumes per year).

(2). The format of each volume is such that each subscriber may tear down the assembled volume into its component articles.

(3). Manuscripts may be sent to any member of the Editorial Board or may be sent directly to The Journal Press. The lag in publication should not be more than a few weeks.

(4). The author will receive 200 reprints of his paper, and subscribers will receive the accumulated publications bound into half-volume units. The second half of each volume will contain the title page and table of contents for that volume. Each half-volume unit will be mailed to the subscribers at its stated date.

(5). Each author will be charged the manufacturing costs of any zincs, halftone cuts, tables, or equations. Excessive use of special fonts (references, footnotes, quotations, appendices) will also be charged.

(6). The publication charges are as follows:

1 to 4 pages	.....	\$40.00
5 to 8	" .....	80.00
9 to 12	" .....	120.00
13 to 16	" .....	160.00
17 to 20	" .....	200.00
21 to 24	" .....	240.00
25 to 28	" .....	280.00
29 to 32	" .....	320.00
33 to 36	" .....	360.00
37 to 40	" .....	400.00

These charges are subject to adjustment as justified.

(7). The accumulated units will be mailed to subscribers bimonthly: January, March, May, July, September, and November. The units may vary in size to accommodate publication needs.

All back issues are available. See announcement on inside front cover.

THE JOURNAL PRESS

2 Commercial Street

Provincetown, Massachusetts 02657

U. S. A.

# Journal of Genetic Psychology

Founded by G. Stanley Hall in 1891; reorganized by Carl Murchison in 1925

Quarterly, beginning with March. Devoted to developmental and clinical psychology, with briefly reported Replications and Refinements and occasional book reviews. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$26.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

---

# Journal of General Psychology

Founded in 1927 by E. B. Titchener and Carl Murchison

Quarterly, beginning with January. Devoted to experimental, physiological, and comparative psychology, with briefly reported Replications, Refinements, and Comments. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$26.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

---

# Journal of Social Psychology

Founded in 1929 by John Dewey and Carl Murchison

Bimonthly, beginning with February. Devoted to studies of persons in group settings and of culture and personality; special attention to cross-cultural articles and notes, and to briefly reported Replications and Refinements. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Three volumes annually. Per annum, \$39.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

---

*Each author will receive 100 unbound reprints gratis. Additional reprints may be purchased at a nominal cost when the author returns page-proof.*

THE JOURNAL PRESS  
2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.

# Club Rate for Five International Psychological Journals

	PER ANNUM
THE JOURNAL OF GENETIC PSYCHOLOGY . . . . . (Quarterly, two volumes per year)	\$26.00
GENETIC PSYCHOLOGY MONOGRAPHS . . . . . (Quarterly, two volumes per year)	\$26.00
THE JOURNAL OF GENERAL PSYCHOLOGY . . . . . (Quarterly, two volumes per year)	\$26.00
THE JOURNAL OF SOCIAL PSYCHOLOGY . . . . . (Bimonthly, three volumes per year)	\$39.00
THE JOURNAL OF PSYCHOLOGY . . . . . (Bimonthly, three volumes per year)	\$39.00
	<hr/>
	TOTAL \$156.00

TEACHERS AND STUDENTS EDITION, *Restricted Sale* \$50.00

The *Teachers and Students Edition* is open to any individual psychologist, teacher, student, or physician but requires an agreement that the edition will not be offered for sale till after five years from publication and an agreement that the edition will not be substituted for an institutional subscription. Write for our Teachers and Students contract card.

## THE JOURNAL PRESS

2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.

## Genetic Psychology Monographs

Founded in 1925 by Carl Murchison

Quarterly, beginning with February. Devoted to developmental and clinical psychology. Each number one or more complete researches. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$26.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

*NOTICE.* Effective immediately we offer prior publication for certain selected monographs. Charges will be based on the announced rates for *The Journal of Psychology*. Regular monographs will carry no page charge for the first 50 pages, but for all pages above 50 there will be a charge of \$10.00 per page. Manufacturing costs of any zincs, halftone cuts, tables, or equations will be charged to authors, as will excessive use of special fonts. Two hundred reprints of each monograph will be supplied free of charge.

The listing of monographs for 1926 through 1963 will be printed in the first half of each volume of all five journals. Monographs published after 1963 will be listed in the second half of each volume.



# Genetic Psychology Monographs (continued)

VOLUME 69—January-June, 1964

- Occupational and industrial therapy in the treatment of psychiatric patients: A controlled study of efficacy in an intensive-treatment institution—M. HUNDZIAK AND B. PASAMANICK  
An analysis of Carl R. Rogers' theory of personality—M. S. KRAUSE  
Personality aspects of teaching: A predictive study—H. ZIMILEK, B. BIEER, W. RABINOWITZ, AND L. HAY  
Situational and personal antecedents of incipient alienation: An exploratory study—W. H. JARRETT AND A. O. HALLER
- Age changes in children's Mosaic responses from five to ten years—L. B. AMES AND F. L. ILC  
A study of the interview process: The caseworker-client relationship—H. J. MEYER, E. F. BORGATTA, AND D. FANSHIEL  
A study of introversion-extroversion as a personality dimension—R. M. ALLEN, H. M. RICHER, AND R. J. PLOTNIK  
Psychological testing with Northwest Coast Alaskan Eskimos—C. E. PRESTON  
A study in family pathology—M. B. HANS

VOLUME 70—July-December, 1964

- Structure and function—H. GUTMAN  
The Lowenfeld Mosaic Test: Norms for five-to-ten-year-old American public-school children and comparative study of three groups—L. B. AMES, F. L. ILC, AND J. AUGUST  
Social and personality correlates of children's estimates of height—J. P. SHAFFER  
The development of ethnic attitudes in New Zealand school children—G. M. VAUGHAN
- The antecedents of two varieties of vocational interests—L. E. TYLER  
Personality in old age—P. E. SLATER AND H. A. SCARR  
Variables affecting experimenter bias in a group situation—R. ROSENTHAL, C. J. FRIEDMAN, C. A. JOHNSON, K. FORD, T. SCHILL, R. C. WHITE, AND L. L. VIKAN  
Ordinal position and sex role identification—B. G. ROSENBERG AND B. SUTTON-SMITH  
Personality correlates of accuracy of self-appraisal in four-year-old children—PAUL S. WEINER

VOLUME 71—January-June, 1965

- Author index for Genetic Psychology Monographs 1926-1964—L. W. FERGUSON  
A theoretical, developmental, and clinical investigation of embarrassment—J. M. SATTLER  
School readiness as evaluated by Gesell developmental, visual, and projective tests—F. L. ILC, L. B. AMES, AND R. J. APPELL  
Rorschach scales of regression in psychosis—S. NEIGER, A. G. SLEMON, AND D. A. QUINN  
The effects of extra tactile stimulation on a group of institutionalized infants—L. CASLER  
Human-eyelid conditioning: The recent experimental literature—L. E. ROSS AND T. F. HARTMAN
- Social status and schoolroom aggression in third-grade children—R. TOIGO  
Overt and covert homosexual problems among male college students—L. J. BRAATEN AND C. D. DARLING  
Role conceptions of the psychiatric aid—M. P. LAWTON AND A. E. GOLDMAN  
Changes in study of values scores during the four years of college—C. W. HUNTLEY

VOLUME 72—July-December, 1965

- A study of process and method in three-year-old twins and triplets learning to read—W. FOWLER  
Individualized teaching with therapeutic aims: A methodological study—M. KOHN AND M. RUDNICK  
Ability and social adjustment at midlife of persons earlier judged mentally deficient—E. L. MILLER
- Development of incentive values in childhood—S. L. WITRYLO, D. J. TYRRELL, AND L. M. LOWEN  
Longitudinal changes in intolerance and authoritarianism for subjects differing in amount of college education over four years—W. T. PLANT  
An NDEA counseling and guidance institute: Prediction, performance, and follow-up—R. H. DANA  
The influence of language on classification: A theoretical model applied to normal, retarded, and deaf children—H. G. FURTH AND N. A. MILGRAM  
Time estimates as affected by incentive class and motivational level—M. FELIX  
Clinical psychology: An historical survey—V. S. SEXTON

VOLUME 73—January-June, 1966

- Vulnerability in infants and young children: A pilot study—G. HEIDER
- Racial preference and ego-identity among white and Bantu children in the Republic of South Africa—A. J. GREGOR AND D. A. MCPHERSON  
Mother-child interaction in cerebral palsy: Environmental and psychosocial obstacles to cognitive development—E. SHERE AND R. KASTENBAUM  
Children's stories—L. B. AMES

VOLUME 74—July-December, 1966

- Changes in personality for groups completing different amounts of college over two years—W. T. PLANT AND C. W. TELFORD  
Brain dynamism in perception of reversible figures—G. K. YACOBZYNSKI  
Changes in Rorschach response throughout the human life span—L. B. AMES  
Mnemonic virtuosity: A study of chess players—A. BINET (Translated by L. M. SIMMEL AND S. B. BARRON)
- The development of behavior patterns in children—B. B. TYLER, F. B. TYLER AND J. E. RAFFERTY  
Correlates of classroom functioning—L. M. SMITH AND B. B. HUGHES  
Children's behavior syndromes and parental responses—R. L. JENKINS, E. NUREDDIN, AND I. SHAPIRO

VOLUME 75—January-June, 1967

- Studies of an elementary game of strategy—B. SUTTON-SMITH AND J. M. ROBERTS  
Child care practices antedating three patterns of preschool behavior—D. BAUMBERG  
The impact of cognitive maturity on the development of sex-role attitudes in the years 4 to 8—L. KOHLBERG AND E. ZIGLER
- Differential factor structures in mathematical ability—P. S. VERY  
Perception of the vertical as a function of affective stimuli and extraneous stimulation in schizophrenic and normal subjects—J. O. WELLS AND W. E. CALDWELL  
Mid-life attainment of the mentally retarded: A longitudinal study—W. R. BALLER, D. C. CHARLES, AND E. L. MILLER

VOLUME 76—July-December, 1967

- Scalogram analysis of some developmental sequences of infant behavior as measured by the Bayley Infant Scale of Mental Development—R. KOHN-RAZ  
The leader and the lost: A case study of indigenous leadership in a poverty program community action committee—L. A. ZURCHER  
A developmental study of the perception of people—R. N. WOLMAN  
Human fears: A factor analytic study of three age levels—G. W. RUSSELL
- Correlates of voluntary social participation—J. E. TEELER  
Personality characteristics and developmental history of creative college women—R. NELSON  
A systematic study of therapeutic interventions in nursing homes—B. A. STOTSKY

VOLUME 77—January-June, 1968

- The fulfillment of promise: 40-year follow-up of the Terman Gifted Group—M. H. ODEN  
A scale for parents' ratings: Some ipsative and normative correlations—N. WALKER  
A methodological investigation of Piaget's concept of conservation of substance—A. C. BITTNER AND M. M. SHINELENG
- Sex differences in mental and behavioral traits—J. E. GARAI AND A. SCHEINFELD

# Genetic Psychology Monographs (continued)

VOLUME 78—July-December, 1968

1. A theoretical and developmental investigation of delayed speech feedback—J. E. WATERS
- Growth changes in factorial structure by age and sex—N. W. DYE AND P. S. VERY
- A multispect investigation of masculine development in kindergarten age boys—H. B. BILDER
- The optimum limits of testing on the Wechsler Intelligence Scales—M. Y. QUEKESHI
- A developmental study of the ego-ideal—L. VAN DEN DAKLE
- The relationship between physical condition immediately after birth and mental and motor performance at age four—N. EDWARDS

VOLUME 79—January-June, 1969

- Adolescents: From the views of the members of an informal adolescent group—T. P. JOSEPH
- Acculturation and childhood accidents among Chinese and Japanese Americans—M. KUSOKAWA
- Dimensions of achievement-related behavior among lower-class Negro parents—D. SOLOMON, R. J. PARELIUS, AND T. V. BUSSE
- Dimensions in the judgment of illness—J. S. KEMRICK
- Reversal shifts and problem-solving abilities at ages 4-7—S. B. KASTNER
- The judgment of body landmark heights—H. NASH
- Recognition of body surface regions—H. NASH

VOLUME 80—July-December, 1969

1. Animal aggression: Review of research—W. M. VERNON
- The effects of positive and negative reinforcement and test anxiety on the reading performance of male elementary school children—S. COTLER
- Personality and psychosomatic illness: A selective review of research—P. R. ROBBINS
- Task and motivational influences on the assessment of creative and intellectual ability in children—N. KOGAN AND F. T. MUMFORD
- A study of an interview method for identifying family authority patterns—R. W. CHAMBERLIN
- Conditioned reinforcing and aversive aspects of the stimuli defining the components of a two-component chain—A. KAUFMAN AND A. BARON
- The Henry Hittleson Center Family Interaction Scales—M. L. BERNHARD, D. I. MEYERS, W. GOLDFARB, N. GOLDFARB, AND N. D. FELDSTEIN

VOLUME 81—January-June, 1970

1. Attitudes of college students and their changes: A 37-year study—V. JONES
- Tattooing among male and female offenders of different ages in different types of institutions—A. J. W. TAYLOR
- Sex differences in mental health—J. E. CARAI
- A selective review of dominance-subordination relations in animals—D. VAN KEEVELD
- The effects of visual or auditory stimulation given in infancy upon subsequent preference for stimulation of both modalities—W. R. BERNKOWITZ
- The influence of certain child-rearing factors upon the development of a structured and salient sense of time—D. ZEAN

VOLUME 82—July-December, 1970

1. Child care attitudes and emotional disturbance among mothers of young children—B. J. COHEN, J. L. WEISS, AND H. U. GRUNBAUM
- A study of intragroup and intergroup relations among experimentally produced small groups—L. N. DIAZ
- The development of role knowledge—D. R. HEISE AND E. P. M. ROBERTS
- Middlecence: Age related stress periods during adult years—J. E. HORROCKS AND M. C. MUSSMAN
- The first six years of life: A report on current research and educational practice—E. LACROIX, P. C. LEE (Ed.), F. LITMAN, D. M. OCLIVIE, S. S. STODOLSKY, AND B. L. WHITE

VOLUME 83—January-June, 1971

1. Measurement of parents' perceptions of their children's development—W. WOLFENBUCHER AND R. A. KUNZE
- The concomitant development of cognitive and moral modes of thought: A test of selected deductions from Piaget's theory—L. C. LEE
- Children's influences upon parental behavior: An attempt to define the relationship with the use of laboratory tasks—J. D. OSOFSKY
- Parent behavior and child academic achievement, achievement striving, and related personality characteristics—D. SOLOMON, K. A. HOULIHAN, T. V. BUSSE, AND R. J. PARELIUS
- Memory factors and school success of average and special groups of ninth-grade boys—M. MEEKER AND C. E. MEYERS
- Personality correlates of vocational planning—R. BARNETT

VOLUME 84—July-December, 1971

1. Perceptual distortion as a function of stimulus objects, sex, naiveté, and trials using a portable model of the Ames Distorted Room—T. A. GERACE AND W. E. CALDWELL
- Galvanotactic behavior and reinforcement of fish *Brachydanio rerio*—W. F. FLANIGAN, JR., AND W. E. CALDWELL
- Self-actualization, types of personality, and neuroses and psychoses—R. DREISTADT
- Children's awareness of their bodily appearance: A developmental study of factors associated with the body percept—E. CLEHNET, J. S. GIGOUS, AND J. COHEN
- Biomegetic phenomena: Some implications for the behavioral and neurophysiological sciences—F. RUSSO AND W. E. CALDWELL
- Sociometric task leadership and the perception of deviance and subsequent conformity, as functions of personality—W. L. LEEBY, JR.
- Murderers who plead insanity: A descriptive factor analytic study of personality, social, and history variables—M. W. KAHN

VOLUME 85—January-June, 1972

1. Children's understanding of negation as a logical operation—S. S. FELDMAN
- Selective auditory perception of parental voices by the disturbed child—S. SANDERS
- Identity type, sociometric ratings, and group behavior—C. N. ZIMET AND C. SCHNEIDER
- Patterns of perceived reaction to frustration among adolescents—J. O. PALMER
- The need for influence: Development and validation of a measure, and comparison with the need for power—J. S. ULLMAN
- C. Lloyd Morgan, E. C. Tolman, and comparative method—E. NEWBURY
- Relearning motor skills in hemiplegia: Error analysis—L. DILLER, J. BUXBAUM, AND S. CHOTELIS
- Effects of participation in rule formation on the moral judgment of children—R. L. MERCHANT AND F. REBELSKY
- Discrepancy and attention in the five-month infant—C. M. SUPER, J. KAGAN, F. J. MORRISON, M. M. HATHE, AND J. WEIFFENBACH

# Genetic Psychology Monographs (continued)

VOLUME 86—July-December, 1972

1. The role of enduring orientations to the environment in personality development—W. C. BRONSON  
Is there an error of the standard? A critique of Piaget's perceptual theory—D. A. BEGELMAN AND G. A. STEINFELD  
A cross-cultural study of the relationship between resistance to temptation and maternal communication—E. R. GREENGLASS  
The intellectual and achievement effects of preschool cognitive stimulation of poverty Mexican-American children—W. T. PLANT AND M. L. SOUTHERN
2. Factors affecting agonistic behavior in several subhuman species—R. F. GOTTIER  
Differential behavioral characteristics of nonpromoted children—J. A. CHASE  
An analysis of cognitive-developmental sequences—J. H. FLAVELL

VOLUME 87—January-June, 1973

1. A survey and interpretation of article characteristics: *The Journal of Genetic Psychology* 1945-1969—J. E. HORROCKS AND J. D. HOGAN  
The development of listener adapted communications in grade-school children from different social-class backgrounds—K. T. ALVY  
Locus of control among schizophrenics and other hospitalized psychiatric patients—T. F. CASH AND J. J. STACK  
Interactive styles and social adaptation—C. N. EDWARDS

THE JOURNAL PRESS  
2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.



## PUBLICATION IN THE JOURNAL OF PSYCHOLOGY

*The Journal of Psychology* provides a special service to authors and readers. Articles accepted for publication appear in print with the least possible delay. Although editorial standards are the same as those for the other journals, all correspondence and handling are expedited, and the referees may, at their discretion, permit some latitude in length of treatment with respect to tabular material, discussion, and review of the literature.

To save time, authors are requested to exercise unusual care in preparing their papers. Special attention should be paid to the General Instructions on the back cover of each issue and to the announcement on the first preliminary page.



\$39.00 per annum  
\$18.00 per volume  
Single Numbers \$9.00

BIMONTHLY  
Three volumes per year  
Immediate publication

March, 1973  
Volume 83, Second Half

(Founded in 1935 by Carl Murchison)

# The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a bilateral contract, and may not be resold for a period of five years from the date of publication.

MARCH, 1973

(Manuscripts are printed in the order of final acceptance)

- Heteromodal cueing and auditory-visual interaction: A literature review 173  
BY BRUCE E. KLAYMAN
- Detection and the timing of a heteromodal cue: Auditory-visual interaction 201  
BY BRUCE E. KLAYMAN
- Factor similarity of personality across private and military samples: An analysis of the Personality/Attitude Schedule . . . . . 215  
BY L. L. CUMMINGS, DONALD L. HARNETT, AND STUART M. SCHMIDT
- Implications for associative processes of switching the middle of the list during serial rote learning . . . . . 227  
BY LINDA L. BEAN AND JAMES H. MCCROSKERY
- Feminism and political radicalism . . . . . 237  
BY MARGUERITE GILBERT FOWLER, ROBERT L. FOWLER, AND HANI VAN DE RIET

(OVER)

Copyright © 1973 by The Journal Press  
Provincetown, Massachusetts, U. S. A.

All rights reserved

Second-class postage paid at Provincetown, Mass  
and additional mailing offices

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

Personal adjustment of hospital staff and their attitudes about mental illness By JOHN FRACCHIA, CHARLES SHEPPARD, AND SIDNEY MERLIS	243
Motivational correlates of Pacific Islanders in urban environments . . . By GEORGE H. HINES	247
Alienation and identity-role diffusion in late adolescence . . . . . By DOROTHY W. JACKSON	251
Tolerance for the unstable and defensive role adjustment in response to sudden physical disability . . . . . By SHELDON KASTNER	257
A preliminary investigation of obscene language . . . . . By RUSSELL FOOTE AND JACK WOODWARD	263
Speed of motor conflict resolution as related to type of conflict and manifest anxiety . . . . . By JEFFREY C. FRACHER AND KENNETH A. BLICK	277
Conservative attitudes and authoritarian values . . . . . By ANTANAS SUZIEDELIS AND MAURICE LORR	287
Group desensitization of test anxiety in elementary school . . . . . By ARREED F. BARABASZ	295
Training of number conservation in retardates . . . . . By ROY P. LANCASTER AND DONALD L. McMANIS	303
Selection of NaCl solutions by sodium-deprived Mongolian gerbils in Richter- type drinking tests . . . . . By JOSEPH W. CULLEN AND ARTHUR E. HARRIMAN	315
A comparison of response rates in response-terminated and time-terminated experimental sessions . . . . . By ALAN C. REPP, SAMUEL M. DEITZ, AND WILLIAM D. WOLKING	323
Selective exposure: An addendum . . . . . By LEONARD L. ROSENBAUM AND ELLIOTT MCGINNIES	329
Language meaning (gender shaping) among blind and sighted students . . . By JOHN G. CULL AND RICHARD E. HARDY	333
Effects of genetic counseling on parental self-concepts . . . . . By MARY ANN ANTLEY, RAY M. ANTLEY, AND LAWRENCE C. HARTLAGE	335

## HETEROMODAL CUEING AND AUDITORY-VISUAL INTERACTION: A LITERATURE REVIEW\*

*New York State Psychiatric Institute*

---

BRUCE E. KLAYMAN

---

### SUMMARY

Experiments were reviewed which dealt with the effect of an auditory cue stimulus on sensitivity to a visual detection stimulus, and *vice versa*. In most cases the heteromodal cue was temporally varied around the time of occurrence of the detection stimulus.

Different results were reported when a prolonged cue was used (which overlapped the detection stimulus when the onsets of both were not simultaneous) than when a brief-duration cue was used: e.g., in the former case, but not in the latter, sensitivity to the detection stimulus was affected by the intensity of a clearly perceptible heteromodal cue. *Es* tended to explain the effects of prolonged cue in "neurophysiological-interaction" terms and of brief-duration cue in "attentional" terms. *Delta t* (difference between times of detection-stimulus and cue onset) was most likely to affect sensitivity to the detection stimulus when marker stimuli reliably informed *S* when to attend for the detection stimulus, but did not exactly define its time of possible occurrence. A consistent *delta-t* effect was that a light-cued sound stimulus was best detected when the flash preceded it by 0 to 500 msec.

---

### A. INTRODUCTION

The present review is concerned with changes in the detectability of auditory or visual stimuli brought about by so-called auxiliary accessory stimuli, or cues. It is often assumed that in *ipsimodal* cueing (i.e., when both the cue and the detection stimulus are delivered in the same sense modality) opportunities for interaction are present at or near the peripheral receptors, or between afferent neural channels over which the stimulus effects are transmitted to the brain, whereas in *heteromodal* cueing (i.e., when the sense modality of the cue differs from the sense modality of the detection stimulus) opportunities

---

\* Recommended by Joseph Zubin of the Editorial Board, received in the Editorial Office on November 27, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

for peripheral or afferent-channel interactions are minimal or absent. Thus, studies of heteromodal cueing may more directly reveal properties of interactions that take place within centers of the brain than studies of ipsimodal cueing. In the present paper are summarized psychophysical auditory-visual experiments in heteromodal cueing, in most of which the timing of the cue-stimulus was systematically varied with respect to the detection stimulus. In order to facilitate the present review, terminology and conventions will be used as described in the next section.

## B. TERMINOLOGY AND CONVENTIONS

### 1. *Markers, SOI, MDI*

In detection trials, stimuli signalling the beginning and/or end of a presentation may be called *markers*, and the detection stimulus is often presented between two markers. However, over trials the detection stimulus may occupy different temporal positions relative to the markers, and it may have an appreciable duration. Thus, the range of time that may be occupied by the detection stimulus will be referred to as the *signal-occurrence interval (SOI)*. An interval between a pair of markers will be defined as a *marker-delimited interval (MDI)* when the pair of markers encloses an SOI, and when, over a block of trials or a session, the SOI enclosed by the markers has a fixed temporal relationship to each of the markers.

The MDI ideally begins with a marker stimulus that instructs the subject to begin attending and ends with a marker stimulus that instructs him to cease attending. In other cases, in which the MDI is less well-defined, either the beginning or the end of the MDI, but not both, may be marked. In these instances *S* may know when to begin attending but not know exactly when to stop, or the reverse may be true. Either of the pair of marker stimuli may or may not be provided by the onset or offset of a cue, and it is possible that more than one marker stimulus may occur simultaneously. The SOI may be smaller than the MDI and contained within it, or equal in duration and coextensive. The SOI may exist in the absence of an MDI, in a trial in which the temporal position of the SOI is unmarked.

### 2. *Negative, Zero, and Positive Delta ts*

In general, the cue stimulus also occurs between the markers; however, the timing of the cue is expressed relative to the time of onset of the detection stimulus. By convention, the time of onset on the detection stimulus is taken as time zero, and when the onsets of the cue and of the detection stimulus are

simultaneous, one speaks of zero time difference or  $\delta t = 0$ . When the cue onset precedes the onset of the detection stimulus, the time difference  $\delta t$  is negative; and when the onset of the cue follows the onset of the detection stimulus,  $\delta t$  is positive. In experiments in which the timing of the cue is variable, one speaks of the range (and distribution) of  $\delta ts$ .

### C. REVIEW OF THE LITERATURE

Experiments in heteromodal cueing of auditory or visual detection stimuli date back at least to the late nineteenth century (5, 20). However experiments in which  $\delta t$  is systematically varied are of much more recent origin: the earliest accessible to this writer was published in 1923, the next in 1937.

#### 1. *Experiments with Simultaneous Heteromodal Cue and Detection Stimulus*

In early studies the heteromodal cue and the detection stimulus were generally simultaneous; prolonged heteromodal stimulation lasting seconds or minutes was delivered, and  $S$ 's threshold to the signal was obtained at intervals in the course of the heteromodal stimulation and sometimes afterward. There was no trial to trial variation in  $\delta t$ , and the temporal characteristics of the two stimuli and of the relations between them were not clearly defined or described.<sup>1</sup> Studies of this nature which were published in western languages are reviewed by Ryan (20) and Gilbert (5); those published in Russian from 1930, by London (14).

Ryan (20, p. 667) concluded that what he called "dynamogenic" studies (intersensory experiments using psychophysical methods) yielded small but reliable effects of heteromodal stimulation on sensitivity. However, he cautioned that the weak mutual influences of abstracted stimulus characteristics are not fairly representative of the much more powerful heteromodal effects involved in the organization of "objects, scenes, events and situations" (p. 690). Gilbert (5), reviewing experiments of the "dynamogenic" type, involving mostly auditory and visual stimuli, concluded (p. 391) that "under conditions of momentary heteromodal stimulation" a relatively intense heteromodal cue will decrease sensitivity to a simultaneous detection stimulus but increase sensitivity to a detection stimulus occurring about  $\frac{1}{2}$  sec later, while a less intense heteromodal cue will increase sensitivity to a simultaneous detection stimulus. Gilbert further suggested that in the case of a "prolonged" heteromodal stimulus its "quality" may determine whether it has an inhibitory or an excitatory effect.

According to London (14, pp. 538-539) Russian researchers, also, reported

<sup>1</sup> A typical study of this type is Kravkov (12).



a "rule of inversion"—increasing heteromodal-stimulus intensity leads "frequently to effects that are the reverse of those induced by weaker intensities." However, doubts were expressed in the Russian literature about whether the "rule" applies to the combination of heteromodal auditory stimulation and visual detection.

London cites reports according to which a simultaneous heteromodal stimulus of moderate intensity decreases absolute sensitivity to brightness in the retinal periphery, while increasing it in the foveal region. In the central retina, sensitivity for the perception of hue ("color sensitivity") is increased for blue-green stimuli, but lowered for orange-red stimuli [see also Kravkov (12)]. (Color sensitivity to the spectral extremes and to light of the yellow region is unaffected by simultaneous auditory stimulation.) These heteromodal effects reach a maximum after the heteromodal stimulus has been on for some time, and may fall off afterward as heteromodal stimulation continues. At least in the case of peripheral brightness sensitivity the effect may reverse after the heteromodal stimulus has been turned off.<sup>2</sup>

The Russians have reported, according to London, that the effect of simultaneous visual heteromodal stimulation on auditory sensitivity varies according to the color of the visual stimulus.

Finally, in a recent study, Bothe and Marks (1) performed a confidence-rating experiment in which the detection stimulus—a 100-msec noise burst—was randomly present in half the trials. The detection stimulus occurred 1050 msec after the offset of a "small, dim" 250-msec warning light. The heteromodal cue was a 200-msec light flash— $2^{\circ}12'$  in visual angle and either 1800, 18, or  $1.8 \text{ c/m}^2$  in intensity—which, when it occurred at all, overlapped the noise burst, so that the noise burst was located in the middle of the MDI which the light flash defined. In the control conditions both the warning light and the noise burst were presented, without the light cue, in either a lightened or a darkened chamber. A signal-detection analysis revealed no difference among the various conditions that was consistent for the four subjects.

## 2. *Variable-delta-t Experiments: Early Studies*

Newhall (17) employed as detection stimulus a 100-msec light flash viewed with the right eye, superimposed upon a fixation cross and adapting field

---

<sup>2</sup> A non-Russian study which found a periodic threshold response to a prolonged simultaneous heteromodal stimulus is Hennebert (6). Here the detection stimulus was a tone, and the heteromodal stimulus was a uniquely unpleasant frequency of flashing light. The absolute auditory threshold oscillated in about 2-min cycles, while tending to move from a high point of sensitivity above the level in the absence of light, reached after about 1-2 min of heteromodal stimulation, to an eventual plateau below the no-light level.

viewed with the left eye. His heteromodal cue was a binaural click. In one experiment Newhall presented his four subjects with *delta t*s of —9000, —4000, and —1000 msec. Trials under each *delta-t* condition were presented in a block; in randomly half the trials the detection stimulus was absent (blank trials). The beginning of the MDI was delimited by the click, but the end was not delimited (except, perhaps, in an approximate way, by *E*'s instruction to "report"). It was found that both the proportion of detection stimuli reported as seen, and the estimated brightness of these suprathreshold flashes, showed some tendency to decrease as *delta t* became increasingly negative, while the false-positive rate did not appear to be affected by the *delta t*. Unfortunately, the interpretation of the data was complicated by the fact that the different *delta-t* procedures were not identical in other ways: In the —4000 and —1000 msec conditions the heteromodal cue was the last click in a five-click series, and under each of the three conditions *S* was given special instructions designed to improve time-estimation.

In a second experiment Newhall delivered a series of threshold flashes, simultaneously accompanied by clicks, and followed it with an identical series of flashes, unaccompanied by heteromodal stimuli. Three of the four *S*s showed a significantly higher detection rate when the visual stimuli were accompanied by clicks (the fourth, no change). The same three subjects produced a larger average estimated brightness of the detected flashes when click was present than when it was absent.

Newhall concluded (p. 74) that more favorable "attentional conditions" increase both the number of stimuli judged suprathreshold and their perceived intensity: "the height of the limen is assumed to vary in some inverse relation to the degree of attention."

Kuroki (13) tested five subjects in an experiment in which a standard and a variable binaural tone were presented on every trial, the variable tone apparently being adjusted from trial to trial to find a point of subjective equality. The tones (474 or 800 Hz) were 1 sec in duration and separated by an interval of 1 sec. A light flash accompanied either the standard or the variable tone in varying temporal relationship. When, at subjective equality, the variable tone was set at a lower intensity when the light accompanied it than when the light accompanied the standard, the flash was assumed to increase the perceived loudness of a concomitant tone.

On the average, all heteromodal conditions increased the subjective intensity of the tone accompanied by the flash. However, a 1/3-sec-duration flash that began simultaneously with the tone (*delta t* = 0) was more effective than either a 2- or 3-sec flash that began 1 sec before the tone (*delta t* = —1000).

On the whole the 2-sec-duration, —1000-msec-*delta-t* flash (which ended simultaneously with the tone) was more effective than the 3-sec-duration, —1000-msec-*delta-t* flash. Also, a 1/5-sec-duration, 0-*delta-t* tone was less effective than the 1/3-sec-duration, 0-*delta-t* tone. *E* did not present any indication of significance in his English summary.

Child and Wendt (2), testing 11 subjects in a Constant-Stimuli experiment, presented a 165-msec-duration, 1000-Hz, right-ear tone as the detection stimulus, and the increase in brightness of a binocular fixation field as the heteromodal cue. *Delta ts* of —2000, —1000, —500, 0, and +500 msec were randomly presented. Light-only and tone-only trials were also presented—whether randomly or in blocks is not indicated. The MDI was delimited by two clicks, 11.2 sec apart, and the detection stimulus occurred always 7.6 sec after the first marker. As *delta t* became increasingly positive (varying from —2000 to +500 msec), percent report of tone, averaged across *Ss*, was found to increase continuously to a maximum at *delta t* = —500 msec, and then to decrease continuously at a rate more rapid than that of the previous increase. Average percent detection under the conditions of *delta t* = —2000 and *delta t* = +500 was higher than under the tone-only conditions. The differences between most conditions were reliable. The authors attributed their results to either “summation of irradiated excitations in the central nervous system” or, less likely, to “relative degrees of alertness produced by the different stimulus-situations” (2, p. 124).

### 3. Maruyama

In a series of parametric studies conducted on six subjects, Maruyama (15) employed as the detection stimulus a 280-msec-duration, 50'-visual-angle flash to the left eye, and as the heteromodal stimulus a binaural tone terminated manually by *E* immediately after the end of the light pulse. *Delta ts* ranged from —5000 to —143 msec; auditory frequencies from 100 to 9000 Hz; loudness levels from 50 to 100 phons; and the retinal location of the test patch from 0° to 40°, along the horizontal meridian of the left eye, from the fovea in the nasal direction. A Method of Limits technique was used: 10 min of dark adaptation was followed by a series of no-tone threshold determinations, then by determinations under various *delta ts* and under the no-tone condition, randomized together. The beginning of the MDI was marked by the onset of the tone; the end of the MDI may or may not have been roughly delimited by tone offset, depending on the constancy of the relationship between tone offset and light offset. It should be noted also that during descending threshold deter-

minations—most determinations were of this type—*S* knew exactly when the detection stimulus would occur.

Plotting parametric sets of functions describing percent change in sensitivity—above or below the no-tone level—as a function of *delta t*, Maruyama came to the following conclusions:

1. Within the effective range of *delta ts*, as frequency of tone decreased, the effect of tone changed from “facilitating” to “inhibitory.” From about 1000 Hz to below 100 Hz there was a band of frequencies which contained the “indifferent frequency” for individual *Ss*. The “indifferent frequency” is the point of transition for an individual from facilitating to inhibitory heteromodal effect.

2. The effect of the heteromodal stimulus seemed to increase in proportion to its loudness, both for “facilitating” and “inhibitory” tones.

3. The more peripheral the retinal location of the test patch, the later after tone onset the peak change in visual sensitivity occurred—whether positive or negative. The maximum change in sensitivity for the 0° (foveal) position was near *delta t* = -143 msec, but for the 40° position the maximum occurred more than 2 sec after tone onset. The author wrote (p. 75) that the “effect spread . . . like wave-motion” from the fovea to the periphery, and his judgment appeared to be confirmed by a family of curves showing percent change in visual sensitivity as a function of retinal location (from 0° to 40°) with time after tone onset (*delta t*) as a parameter. Here, as *delta t* increased from -143 msec to -4000 msec, peak sensitivity shifted from near the 0° position toward the periphery, while the functions became wider and flatter.

A second series of experiments was reported by Maruyama (16). In Experiment I, involving four subjects, the detection stimulus was a 280-msec-duration, 50'-visual-angle flash located 20° nasally along the horizontal meridian of the left eye. The heteromodal stimulus was a 4-sec-duration, 1000-Hz, 80 dB re SPL tone, delivered to either ear. A Method of Limits procedure, similar to that of Maruyama (15) was employed, with descending determinations only and with 30 min of dark adaptation preceding each session.

The 4000-msec MDI was marked by the onset and offset of the tone, so that, with the use of descending threshold determinations exclusively, the possible time of occurrence of the detection stimulus was very well specified.

Maruyama found that tone to the left ear only produced the same kind of (percent change in sensitivity *vs. delta t*) function as was obtained by



Maruyama (15) for the same detection stimulus and a binaural tone of the same loudness and frequency. The right-ear tone, on the other hand, had little or no effect. Since the left-ear tone and the comparable binaural tone produced functions of about the same size, with the peak in both cases at about  $-1500$  msec, and since the right-ear tone was ineffective, *E* concluded that "in the case of binaural stimulation either of two ears has influence on the half of visual field" (p. 83).

Eight *Ss* were tested in a second experiment, in which the heteromodal stimulus was a 3000-Hz, 80-phon, 1-sec-duration tone presented to either the right or the left ear, and the detection stimulus was a 1-sec-duration  $1^\circ$ -visual-angle flash, either  $10^\circ$  nasal or  $10^\circ$  temporal along the horizontal meridian of the left eye. The  $\Delta t$  was 0 in all cases, and a Constant-Stimuli method was used. Visual sensitivity was significantly greater for the combination of right ear and left temporal retina, and left ear and left nasal retina, than for the same visual stimuli in the absence of tone. However, there was no significant increase in sensitivity for the combinations of right ear and left nasal retina, and left ear and left temporal retina.

When the experiment was repeated on seven previously tested subjects, with a flash located  $10^\circ$  temporally along the horizontal meridian of the right eye as the detection stimulus, and a 1000-Hz tone as the heteromodal stimulus, it was found that the left ear, right temporal retina combination produced a significant increase in visual sensitivity; while the right ear, right temporal retina combination produced no change.

In a fourth experiment, for which the subject was *E*, the procedure of Experiment I was repeated with the left-eye target either at the fovea or at one of 10 peripheral positions. The peripheral locations were arranged on a circle whose center was the fovea and whose radius was an angular distance of  $20^\circ$ . Five of the locations were at the nasal half of the retina, and five were at the temporal half. The tone was delivered to the left ear when peripheral-vision sensitivity was being tested, but to either the right or the left ear when testing with the foveal stimulus. Maruyama found that the tone to the left ear produced a strong "facilitating" effect for each of the five positions on the nasal half of the left retina, but the left-ear tone hardly affected the sensitivity of the temporal half. All five nasal-retina curves had a peak at about  $\Delta t = -1500$  msec and, in general, the same size and shape as curves previously obtained for the combination of left-ear tone and a stimulus patch located  $20^\circ$  nasally along the horizontal meridian of the left eye. The foveal curves were facilitating, and peaked near



$\Delta t = 0$ , as before. They were about the same size and shape whether a right-ear or a left-ear tone was used.

It was thus concluded that, as far as peripheral vision is concerned, the right retinal field (left visual field) is influenced only by stimuli to the left ear, and the left retinal field (right visual field) is influenced only by stimuli to the right ear. Foveal vision is influenced equally by stimuli to both ears.

In a final experiment on seven subjects Maruyama investigated whether the "contralateral relationship" between retinal field and ear applied also to the effect of a visual heteromodal stimulus on auditory detection. The detection stimulus was a click to either the right or the left ear, and the heteromodal stimulus was a 200-msec-duration,  $2^\circ$ -visual-angle, 100-milam flash into the left eye, at a location on the horizontal meridian  $20^\circ$  from the fovea in the nasal direction. A Method of Limits procedure was used, and  $\Delta t$  was equal to  $-100$  msec on all trials.

It was found that sensitivity to the left-ear click was significantly greater than under the no-light condition, but that sensitivity to the right-ear click was unchanged. Thus the same laterality relationship was shown to exist for the effect of visual stimulation on auditory sensitivity as for the effect of auditory stimulation on visual sensitivity.

#### 4. Treisman and Howarth

In a series of experiments (8, 9, 22) the detection stimulus employed was a transitory increase in the intensity of a continuous, binaural 500-Hz tone, 60 to 70 dB re SL. The duration of the increment varied between 30 and 300 msec, depending on the experiment, but in most cases was no more than 50 msec. The heteromodal cue was the lighting-up of an unrestrictedly viewed neon bulb, which stayed on until after *S* had responded. There was also a second neon bulb, located above the one providing the cue, which in most cases supplied a "prewarning" stimulus. The interval between the onset of the prewarning light and the first experimental stimulus to occur (either the cue or the detection stimulus) was called the "prewarning interval." In various experiments the prewarning interval had a fixed value in the range of from 1 to 8 sec., or else it varied randomly between 2 and 7 sec. In the auditory-detection experiments reported in these three papers a total of 14 groups were tested, the number of subjects per group (where reported) ranging from two to nine. The experimenters were included among the subjects.

For most groups one of two procedures was used: (a) a Method of Limits

procedure (descending series only) in which determinations were obtained in order of increasing or decreasing magnitude of *delta t*, or (b) a Constant-Stimuli procedure in which the various *delta-t* conditions and the no-cue condition were separated into blocks—one stimulus condition per block—and in which a constant proportion (either 1/5 or 1/11) of trials in each block were blank (no detection stimulus) trials. The range of *delta ts* used in these experiments was from -9000 msec to +1500 msec, but no single group was tested on more than six *delta-t* conditions.

It should be noted that in both types of procedure the subject would probably have been able to judge the time of occurrence of the detection stimulus, in most trials in which cue was present, even when the MDI was not well-defined. This was because a given temporal relation between cue and detection stimulus was delivered a number of times in succession.

Experiments of the above type produced results that may be summarized as follows: The lowest threshold was found when *delta t* = 0; over the range investigated threshold rose continuously as cue onset came increasingly either to precede or to follow detection-stimulus onset. There was no clear-cut effect of prewarning-interval condition, and false-positive rate did not change consistently with change in threshold level of *delta t*. Variability of threshold (measured by standard deviation or standard error) sometimes appeared to decrease with increase in threshold intensity (or in magnitude of *delta t*), but in other experiments no consistent change was found. In the Constant-Stimuli experiments no-cue trials delivered in blocks produced a higher threshold than either *delta t* = -1500 msec or *delta t* = +1500 msec trials delivered in blocks.

These results were essentially the same as the results of experiments (9) in which the detection stimulus was a phosphene-producing 100-msec square electrical pulse applied to the forehead and hand, and the cue was a flash of white or blue light or the ringing of a bell. In these phosphene-detection experiments one of the experimenters' standard auditory-detection procedures was used: descending Method of Limits, with the stimulus conditions delivered in order of increasing negativity of *delta t*.

When a Constant-Stimuli phosphene-detection experiment was performed, with the *delta-t* conditions randomized together, the effect of *delta t* disappeared. However in this experiment there appears to have been no MDI—thus the possible time of occurrence of the detection stimulus was not defined.

In another Constant-Stimuli experiment of this series in which the procedure was nonstandard (22) three stimulus conditions were presented: *delta t* = 0, *delta t* = +500 msec, and no cue. Either the no-cue and +500

msec conditions were presented in blocks, as in the standard-type experiments ("unmixed series"), or the no-cue and +500, and the no-cue and 0, conditions were randomized together ("mixed series"). Under both conditions there was apparently a prewarning interval which varied randomly between 2 and 7 sec. Results showed that there was no significant difference between the mixed no-cue and the unmixed no-cue conditions, between the mixed and unmixed +500 conditions, or between the mixed +500 and the mixed 0 conditions.

In the unmixed +500 condition the subjects knew, as usual, the possible time of signal occurrence (since the same *delta t* was repeated many times in succession), even though the MDI was not well-defined (in this case, only the end of the MDI was marked). In the mixed conditions, the subject had equal knowledge on trials in which the cue was presented. Thus there are several readily apparent reasons why randomization may have had an effect in the phophene-detection experiment (9) but not in the auditory-detection experiment which has just been discussed. Apart from the sketchiness of the data of the present experiment, there are differences in stimulus modality and—perhaps more important—in the extent to which the possible time of occurrence of the detection stimulus was defined, under the randomized procedure.

Howarth and Treisman (10) tested three research psychologists (including one of the *Es*) in a Constant-Stimuli experiment in which the detection stimulus was a 40-msec increment in the intensity of a constant binaural 500-Hz tone, 60-70 dB re SL, and the heteromodal cue was a square spot of light, 14' in visual angle and about 10 msec in duration, focused on a fixation cross. The viewing of the heteromodal cue was presumably binocular. There was a "prewarning" light located to the right of the stimulus spot, and the "prewarning interval" (defined above) varied randomly between 2 and 4 sec. Two near-threshold intensity levels of the heteromodal cue were employed—"Strong," detected about 95% of the time, and "Weak," detected about 50% of the time—and there were two values of *delta t*—0 and —200 msec. The two *delta-t* conditions were presented in separate blocks; in each block the heteromodal cue was randomly and equally often "Strong," "Weak," or absent. Also, 1/5 of the trials in each block were randomly no-detection-stimulus trials. There was a series of light-detection trials—half of the trials with the Strong and half with the Weak stimulus—presented at the beginning and the end of every session.

It should be noted that in this experiment the MDI was badly defined: only the beginning of the MDI was marked, by the heteromodal cue, and only in

trials in which the cue occurred and was seen. However, judgment of the possible time of occurrence of the detection stimulus was aided by the presentation of the two *delta-t* conditions in separate series and by the fact that "the subject was always told before each series what time interval to expect between the two stimuli and was given an initial trial at the beginning of each series with high values of both stimuli . . ." (10, p. 13).

In describing the results for the *delta t* = 0 series the *Es* noted that in their earlier experiments heteromodal cues well above threshold produced a maximum lowering of auditory threshold of about .1 dB. The "Strong" visual stimulus in the present experiment—detected 94% of the time overall when it was presented alone—produced a lowering of auditory threshold of .074 dB, compared with the threshold for the no-cue trials with which the *delta t* = 0 trials were randomized. The "Weak" visual stimulus produced a lowering of auditory threshold of only .034 dB, but since it was detected only 48% of the time, the difference could be accounted for by the assumption that the heteromodal stimulus had "a full strength effect provided it were seen" (p. 15).

In the case of the *delta t* = -200 msec series, the results were somewhat different: there was no difference between the effect of the Weak and the Strong heteromodal cues, and the lowering of threshold accounted for by either cue was only about .05 dB. Also, there was no difference between thresholds obtained at *delta t* = 0 and *delta t* = -200 msec when the Weak cue was used, although for the Strong cue the difference was significant in the expected direction.

The authors concluded (p. 16): "The fact that the intensity of the warning signal [heteromodal cue] has so little effect is evidence for the theory that the warning acts simply as a 'marker' [of the temporal position of the detection stimulus] . . . The effects of warning in all our experiments are extremely small, of the same order as the variability of the threshold measurements."

Treisman (21) reported two additional Constant-Stimuli experiments. In the first, five subjects were tested on the detection of a 40-msec increment in the standard tone (see above). The cue was a 40-msec-duration, binocularly viewed flash—illuminating a 4°40' fixation circle—which could take either of two intensities—"Strong" (180 ft-lam) or "Weak" (.56 ft-lam). The brightness of the unilluminated stimulus patch was .40 ft-lam, and the brightness of the surround was 3.2 ft-lam. There were three cue-presentation conditions, presented separately in blocks: *delta t* = -1500 msec, *delta t* = 0, and no cue. No-detection-stimulus trials constituted randomly 1/5 of the



trials in each block. The prewarning interval varied randomly between 2 and 5 sec.

In this experiment the beginning of the MDI was not marked for the no-cue condition, and the end of the MDI was not marked for any of the three cue-presentation conditions. However, in the  $-1500$  and  $0$  conditions, estimation of the possible time of detection-stimulus occurrence was aided by the fact that the *delta-t* conditions were presented separately.

Results showed that there was no significant difference between the effect of the two cue intensities on threshold at either *delta t*. As in earlier experiments the various cue-present conditions produced significantly lower average thresholds than the no-cue condition, and differences in threshold level were not related to differences in false-positive rate. However, the  $-1500$  threshold was higher than the  $0$  threshold (significance not stated) for the "Weak" condition only. The author considered that his results tended to refute the theory that the effect of a heteromodal cue on threshold is due to arousal produced by the cue: cues of such different intensities ought to be arousing to different extents and to produce different effects on threshold.

In the second experiment the general stimulus conditions were the same as in Howarth and Treisman (8, 9) and Treisman and Howarth (22). The detection-stimulus increment was 50 msec in duration, and the prewarning interval was 1 sec. The eight subjects were tested in separate "Constant" and "Randomized" sessions. In the Constant sessions the procedure was of the standard type: *delta ts* ( $-2750$ ,  $-2250$ ,  $-1750$ ,  $-1250$ ,  $-750$ , and  $-250$  msec) were presented in separate blocks, one *delta t* per block. In the Randomized sessions "ranges" of *delta ts* were presented in blocks; a given block contained one of the following three ranges:  $-3000$  to  $-2000$  msec,  $-2000$  to  $-1000$  msec, or  $-1000$  to  $0$  msec. *Delta ts* were randomized within each range (rectangular distribution). No-detection-stimulus trials randomly constituted 1/21 of the trials in each block, whether Constant or Randomized. Ss were told the *delta t* or range of *delta ts* to expect before each block of trials. Thus neither the beginning nor end of the MDI was marked in the Randomized sessions, but the beginning of the MDI was marked in the Constant sessions, by the occurrence of the cue, since *delta-t* conditions were presented separately in blocks. Judgment of the possible time of occurrence of the detection stimulus in the Constant sessions was further aided by the fact that S was told what *delta t* to expect before the block began.

Results showed that the Constant sessions produced the expected effect—overall percent detection fell almost continuously as *delta t* increased from



—250 to —2750 msec. For the Randomized sessions the results were somewhat different: percent detection was approximately constant (60-62%) as *delta t* increased from —250 to —1750 msec, then fell abruptly and became constant again (at 50-51%) from —2250 to —2750 msec. The data were specified in terms of class intervals 500 msec apart, so that the detail of the function is unknown. By fitting two straight lines to the data (in a manner not convincing to the present reviewer) *E* concluded that, as *delta t* increased, percent detection remained constant until about —1500 msec and then fell continuously. He explained the results, as he interpreted them, on the basis of the Range of Expectation hypothesis.

The Howarth and Treisman (9) and Treisman (21) papers provide the most detailed presentation of Treisman and Howarth's theoretical interpretation of their results. The entire body of their experimental data (except for data concerning positive *delta ts*) is explained by means of the Range of Expectation Hypothesis: On every trial the subject gradually lowers, then raises, his threshold in a range about the expected time of arrival of the detection stimulus. The maximum lowering occurs at the time when the likelihood of signal occurrence is judged to be greatest; throughout the range of expectation the amount of lowering of the threshold at any given point in time is proportional to the judged likelihood of detection-stimulus occurrence at that time. Where a (negative or zero) *delta t* is repeated over successive trials, the time of maximum lowering of threshold occurs, averaging over trials, at about the time the detection stimulus is in fact delivered. However, *S* cannot restrict the lowering of his threshold to the time of signal occurrence because his estimate of that time is somewhat variable, as he has learned. For a given set of stimulus conditions the length of his range of expectation corresponds to the range of variability of his estimates of the time of signal occurrence.

It is thought that the range of expectation can be divided into successive "moments," each moment perhaps 100 msec long. During each moment a sampling process takes place. According to this analysis the lowering, and then raising, of threshold during the range of expectation is actually a lowering and raising of the likelihood-ratio criterion; approximate constancy of the false-positive rate between the various *delta-t* conditions is due to the fact that from moment to moment in the range of expectation the shifting of the likelihood-ratio criterion is compensated for by a change in the actual probability of signal occurrence. If at any moment during the trial the subject detects a "central effect" which exceeds the criterion obtaining during

that moment, the subject announces, at the end of the trial, that he has perceived the detection stimulus.

S is assumed to be able to estimate short intervals more accurately than long ones. Thus, for short *delta ts* his range of expectation is narrower than for long *delta ts*, and for short *delta ts* the maximum lowering of threshold (at the central moment of the range of expectation) is greater than for long *delta ts*. The reason for the latter statement is that lowering of threshold at a given moment covaries with the judged likelihood of signal occurrence at that moment, and whether the range of expectation is small (contains few moments) or large, the sum of the momentary likelihoods of signal occurrence across moments must equal 1. In general then, the larger the negative *delta t*, the longer the range of expectation and the shallower the depression in threshold during that range. It was speculated that this reasoning may explain why, in some experiments, variability of threshold was greater for lower thresholds and smaller *delta ts*: the shorter the range of expectation, the greater the difference between the "normal" threshold level at the beginning and end of the range and the maximally depressed level in the center of the range.

This theoretical structure was used to explain Treisman's interpretation of the results of his second experiment (21). The author concluded that, in the Randomized sessions, when a collection of sufficiently long *delta ts* was delivered, S developed a range of expectation that was considerably longer than the range of randomization, so that randomization had little effect. When a collection of sufficiently short *delta ts* was delivered, the range of randomization was larger than the range of expectation, obliterating the property of temporal specification of the cue. Thus in the latter case percent detection did not change as *delta t* decreased further.

#### 5. Watkins et al.

In these experiments a distinction will be made between an *incremental* and a *decremental* cue, whether the cue is a heteromodal or an ipsimodal one. An incremental cue is provided by a transitory increase in the level of a constant or zero-level stimulus, followed by a return to the original level. A decremental cue is provided by a transitory decrease in the level of a constant stimulus, followed by a return to the original level.

Watkins (23) tested six Ss in an experiment in which the detection stimulus was a  $\frac{1}{2}$ -sec-duration series of five evenly spaced 10-microsec-duration flashes presented on an otherwise darkened, binocularly viewed circular field, about

6' in visual angle. The field was at the center of an illuminated circular surround, 9°45' in visual angle. A temporal forced-choice method was used, with four 500-msec-duration SOIs, with 500-msec separations between them, presented on every trial, and with the duration of the detection stimulus coextensive with the SOI in which it occurred. The four heteromodal conditions were presented separately in 30-trial blocks, and *S* was informed of his "score" (number correct) at the end of each block. The *delta t* was 0 in all cases.

The heteromodal stimulus was filtered binaural white noise, 75 dB re SPL, overall. There were two conditions of (incremental) heteromodal cue: continuous noise presented only during the SOIs, and auditory flutter (10/sec; noise-cycle ratio,  $\frac{1}{2}$ ) presented only during the SOIs. There were also two control conditions: continuous 75 dB re SPL noise presented during the entire block of trials, and 75 dB re SPL auditory flutter presented during the entire block of trials. In all cases there was also a (decremental) ipsimodal cue—panels just above and below the surround darkened during the SOIs. Thus the MDI—which was coextensive with the SOI and with the duration of the detection stimulus, when the detection stimulus occurred—was marked by the onset and offset of the ipsimodal cue on all trials, and also by the onset and offset of the heteromodal cue, when the heteromodal cue occurred.

It was found that noise during the SOI resulted in about 1.2 times as many correct detections as continuous noise; there were no other significant variables and no interactions.

Watkins and Fehrer (26, 27) presented a series of experiments<sup>3</sup> in which the detection stimulus was an approximately 165-msec-duration increase in the brightness of a 25-ft-lam, binocularly viewed, circular fixation spot, 34' in visual angle, which was located in the center of an illuminated square surround. A temporal forced-choice technique was used in which every trial contained four SOIs, each about 165 msec in duration, separated by intervals of about 495 msec. The SOI was coextensive with the duration of the detection stimulus, when the detection stimulus occurred.

The heteromodal cue was an increment or decrement in the intensity of filtered binaural white noise. The duration of the increment or decrement was either 165, 330, or 495 msec, approximately. In the 1964 experiments (26) the incremental condition consisted of the presentation of 70 dB re SPL noise at or near the SOIs and 50 dB re SPL noise elsewhere in the block of trials; the decremental condition consisted of the presentation of 50 dB

<sup>3</sup> For details of the apparatus and stimulus sequences see Watkins, Nickerson, and Schjelderup (29).

noise at or near the SOIs and 70 dB noise elsewhere. In the 1965 experiment (27) the two noise levels used were 70 dB re SPL and 0. The *delta ts* employed were -330, -165, 0, and +165 msec. There were four control conditions: in the 1964 experiments noise was constant at either 50 or 70 dB during the block of trials; in the 1965 experiment noise was either constant at 70 dB or continually off. The various combinations of heteromodal-stimulation conditions were presented separately in blocks.

The MDI was defined in almost every case; it was either 165, 330, or 495 msec in duration, and its beginning or end was marked by one or more of a variety of cues: increase or decrease in noise level, or increase or decrease in the brightness of either one of two ipsimodal stimuli—a 36'-visual-angle "timing spot," concentric with the fixation field; or a "cue panel," located at the lower left corner of the surround. When the MDI was 165 msec in duration, the SOI was coextensive with it, and the SOI occupied one of the successive 165-msec segments within the 330- or 495-msec MDIs.

A total of eight groups were tested in the present experiments; the number of Ss per group ranged from two to seven.

The results of these experiments can be summarized as follows. All other variables being equal: (a) The incremental heteromodal cue produced a higher detection rate than the decremental heteromodal cue, where the change in intensity level in both cases was equal, and where the cues were coextensive with the SOIs. The decremental heteromodal cue produced a higher detection rate than constant heteromodal stimulation at a comparable level. (b) The  $\Delta t = 0$  condition produced a higher detection rate than the negative- $\Delta t$  conditions. The negative  $\Delta ts$  produced higher detection rates than  $\Delta t = +165$ . (c) Heteromodal cues terminating simultaneously with the end of the SOI produced higher detection rates than heteromodal cues terminating after the end of the SOI. (d) When an incremental ipsimodal cue (the brightening of the timing spot) was presented coextensively with the SOIs, the effect of the heteromodal cue was diminished—i.e., the differential effect of incremental, decremental, and constant heteromodal stimulation was reduced.

Watkins (24, 25) and Watkins and Schjelderup (28) tested a total of 12 groups, with two to eight subjects per group, on a series of experiments in which the detection stimulus was a 1011-Hz binaural tone—170, 200, or 250 msec in duration—and the heteromodal stimulus was a binocularly, unrestrictedly viewed 2° circular patch. For the incremental heteromodal cue the brightness of the stimulus patch was, in most experiments, 1350 ft-lam at or near the SOIs, and 4 ft-lam elsewhere in the block of trials; for the de-



cremental heteromodal cue the brightness in most experiments was 4 ft-lam at or near the SOIs, and 1350 ft-lam otherwise. In the constant heteromodal conditions the light was maintained at either 4 or 1350 ft-lam continuously during the block of trials. The ipsimodal stimulus was filtered binaural white noise. For the incremental ipsimodal condition the noise was at 60 dB re SPL during the SOIs and at 40 dB elsewhere in the block of trials; for the decremental ipsimodal condition the noise was at 40 dB during the SOIs and at 60 dB otherwise. The constant ipsimodal condition was continuous 60 dB noise. The ipsimodal cue was always coextensive with the SOI, but *delta t* for the heteromodal cue, although 0 in most cases, was also -500, -250, -200, +200, and +250 msec in various experiments. The duration of the heteromodal cue was generally the same as the duration of the detection stimulus, although in one experiment it could be 80 msec longer, and in another 200 msec longer (28). As before, a temporal forced-choice procedure was used in which each trial contained four SOIs (the duration of each SOI coextensive with the duration of the detection stimulus). The various combinations of heteromodal and ipsimodal cue conditions were presented separately in 48-trial blocks, and *S* was told his "score" at the end of each block.

The MDIs were defined in all cases—by the ipsimodal cue, the heteromodal cue, or both; they were usually coextensive with the SOIs.

The results were essentially as follows: Where ipsimodal stimulation was constant, the differential effect of the heteromodal stimulus was greater for light-cued tone detection than for noise-cued light detection. The differential effect, as stated before, is that an incremental heteromodal stimulus coextensive with the SOI produces a higher detection rate than a decremental heteromodal stimulus, and a decremental heteromodal stimulus produces a higher detection rate than constant or zero-level heteromodal stimulation. However, an ipsimodal cue coextensive with the SOI tended to remove the differential effect. An incremental ipsimodal cue tended to remove the differential effect almost completely; a decremental ipsomodal cue tended to remove the differences only partially. Some of these results are summarized by Watkins (24). Average percent correct for the incremental heteromodal condition minus average percent correct for the decremental heteromodal condition equals, when the detection stimulus is embedded in a constant ipsimodal background, 11.8% for tone detection and light heteromodal cue, and 6.7% for light detection and noise heteromodal cue. When the signal is embedded in an incremental ipsimodal cue (light flash or noise burst), the difference is reduced to .2% for tone detection and light heteromodal cue, and 2.2% for light detection and noise heteromodal cue. In addition when the tone



signal was embedded in a decremental ipsimodal cue (noise reduction), the difference was 3.1% (25).

Added findings (28) showed that *delta ts* of 0 and -200 msec tended to produce the highest detection rates, compared with the other *delta ts* used, and the magnitude of the increment in the heteromodal cue had no effect. Also, confirming earlier results, higher detection rates were found when the heteromodal cue ended simultaneously with the end of the SOI than when the heteromodal cue ended 200 msec after.

The authors offered no comprehensive theoretical explanation of their findings. They merely noted that the arousal or attention-heightening properties of the heteromodal cue seemed more important than its time-specification properties in determining the effect of heteromodal cueing on detection. There were two reasons for this hypothesis:

1. The incremental heteromodal cue was clearly more facilitating than the decremental heteromodal cue, when both cues were coextensive with the SOIs and when the magnitude of the increment in one case equalled the magnitude of the decrement in the other. Yet both kinds of cue conveyed exactly the same temporal information, although they may not have been equally arousing (26, 27). The differential effect of the incremental *vs.* decremental heteromodal cue may be related to "a property of the non-specific 'alerting' or 'activating' system . . . viz., the response is greater to excitatory . . . stimuli than to stimulus-intensity-reducing varieties" (24, p. 37).

2. The presentation of a decremental ipsimodal cue coextensive with the SOI reduced, but failed to eliminate completely, the differential effect of incremental *vs.* decremental heteromodal cue on tone detection. Thus "it is difficult to understand . . . why auditory perceptual mechanisms should rely upon visual information for temporal data, in the presence of a dependable and conspicuous auditory time cue . . ." (25, p. 478).

#### 6. Other Recent Studies

Egan, Schulman, and Greenberg (4) tested eight subjects in an experiment<sup>4</sup> in which the detection stimulus was a 250-msec-duration, 1000-Hz binaural tone, and the heteromodal cue was a 500-msec-duration light. There was constant ipsimodal stimulation: binaural noise, about 65 dB re SPL. *Delta ts* were -1000, -500, -250, 0, +500, +1000, and +2000 msec, and the various *delta-t* conditions were presented separately in blocks. Randomly half the trials in each block were no-detection-stimulus trials. The MDIs were defined only by the onset and offset of the heteromodal cue: thus both the

<sup>4</sup> For details of procedure see Egan, Greenberg, and Schulman (3).

beginning and the end of the MDI was defined for the  $\delta t = 0$  and  $\delta t = -250$  msec conditions only (in both of which the SOI occurred during presentation of the heteromodal cue). For positive  $\delta t$ s the end only of the MDI was defined, by cue onset; for negative  $\delta t$ s greater than  $-250$  msec the beginning only of the MDI was defined, by cue offset. However *S* was given a demonstration of the  $\delta t$  to be presented before every block of trials, with the noise turned off.

After every trial *S* rated his confidence that a signal had occurred. *Ss* had had extensive practice in the use of this technique.

Results were in the form of the "detectability index" ( $d_s$ ), an estimate of  $d'$ . The detectability index was greatest at  $\delta t = -250$  msec, and significantly greater at  $-250$  msec than at 0. The detectability index as a function of  $\delta t$  showed a continuous, upward-concave decline in both directions from  $-250$  msec, the decline being steeper in the positive direction.

The *Es*' explanation of their results is of the time-specification type. At  $\delta t = -250$  msec, the 250-msec-duration SOI overlaps the second half of the 500-msec-duration cue. As negative  $\delta t$  increases beyond  $-250$  msec, *S*'s memory of the temporal location of the SOI becomes less accurate. As positive  $\delta t$  increases from 0, two factors operate: (a) memory of the temporal location of the SOI becomes less accurate, and (b) memory of the "input waveform" becomes less accurate.

Finally, two experiments in heteromodally cued visual detection will be summarized, which have a less direct relationship to the present study than the experiments previously discussed. In one experiment, a single visual detection stimulus was employed, but the heteromodal cue was electrocutaneous; in the other experiment, the heteromodal cue was a click, but the detection stimulus was a pair of flashes, and a two-flash threshold was obtained.

Novak (18, 19) tested two subjects in an experiment (Experiment B) in which the detection stimulus was a monocularly viewed 1-msec flash on a circular foveal test patch  $.80^\circ$  in visual angle, and the heteromodal cue was a 1-sec DC square-wave electrocutaneous pulse to the third finger of the right hand. The electrocutaneous pulse was 6 dB re SL and was not intense enough to produce a reflex twitch or to be reported as aversive. The dark-adapted subjects were tested on a Method of Limits procedure (ascending trials only); there was a series of no-cue determinations at the beginning of every session ("presession control"), and blocks of no-cue determinations were scattered through the rest of the session in a fixed order ("intrasession control"). The approximate range of  $\delta t$ s was from  $-300$  to  $+300$  msec; the determinations were randomized according to  $\delta t$ , within limits. *S* initiated each trial himself. Since in a given Method of Limits determination the same

$\delta t$  is presented several times in succession,  $S$ 's act of initiating the trial may be said to mark the beginning of the MDI; the end of the MDI was apparently not defined.

It was found that, for each  $S$ , threshold luminance was minimal at  $\delta t = -25$  msec, and increased more-or-less continuously as  $\delta t$  varied in both the positive and negative directions. At around  $\delta t = -190$  msec and  $\delta t = +150$  msec threshold luminance reached asymptotes at the intrasession-control level. The experimenter explained the results on the basis that "the value of  $-25$  milliseconds . . . [may represent] a difference in conduction time for impulses produced by the electrocutaneous and visual stimulation to arrive at some locus (or loci) in the central nervous system" (19, p. 15).

It might also be pertinent to mention a study of Horn and Venables (7) in which the effect of a heteromodal cue on two-flash threshold was investigated. The flashes, which were produced by diffuse illumination of a cathode-ray tube, were each 200 ft-lam in brightness and 5 msec in duration. The heteromodal cue was either of two unspecified loudness levels of click, 18 dB apart. The click preceded the first flash of the pair at  $\delta ts$  ranging from 0 to  $-600$  msec. The  $\delta ts$  were presented according to the row orders of a Latin Square. No-cue thresholds were determined at the beginning ("initial threshold") and end ("final threshold") of each session.

The results were presented in terms of the largest separation between the two flashes at which they were reported as one flash. The two groups of eight and 12 subjects produced average two-flash thresholds that were highest at  $\delta t = 0$  and decreased continuously as  $\delta t$  increased in the negative direction. As  $\delta t$  increased, two-flash threshold reached the final-threshold level at about  $\delta t = -150$  msec, and the (lower) initial-threshold level at about  $\delta t = -400$ . Increasing the intensity of the click appeared to increase the heteromodal effect (i.e., further raise the threshold compared with the control levels) at the  $\delta ts$  smaller than  $-150$  msec. An electrocutaneous heteromodal cue (10-msec-duration square-wave pulse delivered to the forearm) produced the same kind of results as the click cue. The authors think that their effects may be related to reticular modification of cellular responses at the cerebral or striate cortex, and may parallel the "psychological refractory period" (7, p. 296).

#### D. CONCLUSIONS

Certain conclusions emerge concerning the effect of the variables of  $\delta t$ , degree of specification of the MDI, and duration of the heteromodal cue. Of these variables the duration of the heteromodal cue is, in a sense, the most

theoretically important, since experimenters who employ a prolonged heteromodal cue, overlapping the detection stimulus, favor a "neurophysiological-interaction" type of explanation of the effect of heteromodal cueing on detection, while experimenters who employ a brief or momentary heteromodal cue tend to favor one of two types of "attentional" explanation—specification of the time of detection-stimulus occurrence (i.e., Treisman, Howarth) or arousal effects of the cue (i.e., Watkins *et al.*). A summary of the effects of these three variables is presented directly below.

### 1. *Effect of Varying Delta t*

Where the set of *delta t*s employed in a given experiment was suitable for determining the *delta t* at which maximum detectability (sensitivity) occurred, the location of the point of maximum detectability was found to be in the range of from  $\text{delta } t = 0$  to  $\text{delta } t = -500$  msec, for light-cued sound-detection experiments (2, 4, 8, 21, 22, 28). This conclusion is in general agreement with the findings of Watkins and Schjelderup (28), who plotted the data of four light-cued sound-detection experiments (all included in the present review) on a composite graph depicting "proportional auditory achievement" as a function of *delta t*. These authors concluded that maximum detectability was in the region of  $\text{delta } t = -300$  msec.

There are not enough light-detection data to yield a clear picture of the situation. However, experiments involving sound-cued detection of a foveal stimulus yielded, in one case, a maximum at  $\text{delta } t = 0$ , when the heteromodal cue did not extend beyond the end of the detection stimulus (26); in another case (15, 16) the maximum was found to be at a *delta t* no more negative than  $-143$  msec. An experiment involving electrocutaneously cued detection of a foveal stimulus (18, 19) yielded a point of maximum detectability at  $\text{delta } t = -25$  msec. For sound-cued detection of peripheral visual stimuli, Maruyama (15, 16) found the maximum to be located at *delta t*s of  $-1000$  msec or more.

The question of the location of the maximum on the detectability-*vs.*-*delta-t* function is complicated by the wide differences in stimulus conditions and procedure between different experiments, as well as by the individual differences that probably exist in this kind of performance. (There is, of course, no reason why there has to be only one maximum.) Another complicating circumstance is that the set of values of *delta t* was different in different experiments, and the separation of *delta-t* values within a given set was generally too large to get a clear idea of the detail of the detectability-*vs.*-*delta-t* function.



As  $\Delta t$  varies from the point of maximum detectability in either the positive or the negative direction, detectability decreases more or less continuously toward an asymptote: the decrease is more rapid in the positive direction (2, 4, 8, 9, 15, 16, 18, 19, 21, 22, 28). Detectability has been reported to continue to decrease beyond  $\Delta t = -4000$  msec in the negative direction (9) and  $\Delta t = +1000$  msec in the positive direction (22).

It is not possible to determine at what  $\Delta t$ s the function reaches the "baseline" level of detectability in the absence of a cue, or even if the maximum detectability when cue is present is consistently higher than no-cue detectability. Experimenters used either of two measures of baseline detectability, which may not have had the same meaning—no-cue trials presented in blocks and no-cue trials randomized together with the cue-present trials—and in some cases did not specify which measure they used, or used none at all.

## 2. Specification of the MDI

The effect of the heteromodal cue appears to be greatest when the SOI is located within an MDI—that is, when it is preceded or followed, but preferably both, by a temporal marker—but when the MDI is larger than the SOI rather than coextensive with it. Two lines of evidence lead to this conclusion:

1. Randomization of  $\Delta t$ s in a Constant-Stimuli procedure was compatible with the obtaining of a detectability-*vs.*- $\Delta t$  function where both the beginning and the end of the MDI were defined (2), but the function was more or less abolished by randomizing  $\Delta t$ s where neither the beginning nor end of the MDI was defined [phosphene-detection experiment (9); also (21)]. Where only the beginning or end of the MDI, but not both, was defined, the results of randomization were inconclusive (22).

2. Presentation of an ipsimodal cue coextensive with the SOI, so that the beginning of the cue marked the beginning of the SOI, and the end of the cue marked the end of the SOI, weakened or obliterated the difference in the effect of incremental and decremental heteromodal cues (24, 25, 26, 27). There was a suggestion that the detectability-*vs.*- $\Delta t$  function was weakened also [comparison of the results of Experiments III and IV (26)].

It is possible that the above results are related to time-specification via the heteromodal cue: A clearly defined MDI, larger than the SOI, restricts the range of possible intervals between the detection stimulus and the cue; under the restriction the cue can affect temporal specification of the detection stimulus. Where there is no MDI, on the other hand, the range of possible



intervals between the detection stimulus and the cue is much less limited, so that *S* would have much more difficulty in acquiring the temporal relationship between cue and detection stimulus. In the third possible case, where the SOI is marked exactly—by a pair of marker stimuli that define an MDI coextensive with the SOI—the time-specification property of the heteromodal cue is redundant [see Discussion in Watkins and Feehrer (26)].

It should be noted that there are other possible aids to estimation of the temporal location of the SOI in the trial besides the MDI. For example, *S* could be told how long after the heteromodal cue the detection stimulus would occur, and be instructed to time-estimate the *delta t* (17), or *S* could be given descending Method of Limits trials (9, 15, 16, 22) so that, in effect, each *delta t* would be demonstrated to him several times before *S* had actually to detect the signal at threshold. Such a condition is logically similar to the situation of *delta-t* conditions being separated into blocks in a Constant-Stimuli or Yes-No experiment, with a demonstration of the stimulus sequence before each block (4, 10). Even in the case of ascending Method of Limits trials, where *delta-t* determinations are randomized (e.g., 15, 18, 19) it seems likely that *S* could obtain some information about the temporal location of the SOI, even when a well-defined MDI was not present, before an actual threshold determination took place.

### 3. Duration of the Heteromodal Stimulus

It appears that these experiments can be classified in one of two categories: (a) experiments in which the heteromodal stimulus is of brief duration, and (b) experiments in which the heteromodal stimulus is prolonged, so that for negative *delta ts* it extends until the end of the SOI, or after. In the first type of experiment the effects of heteromodal cueing appear to be mainly due to the attentional properties of the heteromodal stimulus, properties of either the arousal or time-specification variety, or both; in the second type of experiment the heteromodal effects appear to be mainly due to neurophysiological interaction between the traces of the heteromodal and detection stimuli.

The basis for this line of reasoning is that the results produced by a prolonged heteromodal stimulus appear to depend on its intensity—in either a proportional or a “rule of inversion” relationship, according to the situation (5, 14, 15); but where the heteromodal stimulus is brief, its intensity appears to have no effect, provided the heteromodal cue is intense enough to be perceived all of the time (10, 21, 28). It has also been reported that the effects of heteromodal cueing depend on the frequency-wavelength and

laterality characteristics of both the detection stimulus and the heteromodal stimulus where the heteromodal stimulus is prolonged (14, 15, 16). Presumably these findings, also, would not apply in the case of a heteromodal stimulus of brief duration.

An additional finding is that, in the case of a prolonged heteromodal stimulus, detectability is greater when the heteromodal stimulus ends simultaneously with the end of the SOI, than when it ends after (13, 26, 28).

Thus, prolonged cue appears as functionally different from momentary cue, perhaps because different and more direct physiological interactions become possible.

### E. SUGGESTED FURTHER RESEARCH

It might be useful to differentiate between the effect of sound cue on sound detection, of sound cue on light detection, of light cue on sound detection, and of light cue on light detection—something that could not be done on the basis of the experimental results reviewed in this paper. It is possible that such a comparison would suggest whether the nature of any obtained effects (*a*) depended on the modality of the cue alone, (*b*) depended on the modality of the detection stimulus alone, (*c*) depended on the modality of both the cue and the detection stimulus, or (*d*) was independent of the modality of both the cue and the detection stimulus.

If *proposition a* is true, sound-cued sound detection and sound-cued light detection should produce results similar to each other, as should light-cued sound detection and light-cued light detection, but the effects of the two sound-cued techniques should be dissimilar to the effects of the two light-cued techniques. If *proposition b* is true, the results of sound-cued sound detection and light-cued sound detection should be similar to each other, as should the results of sound-cued light detection and light-cued light detection, but the results of the two sound-detection techniques should be dissimilar to the results of the two light-detection techniques. If *proposition c* is true, the results for sound-cued sound detection, sound-cued light detection, light-cued sound detection, and light-cued light detection should all be dissimilar; while if *proposition d* is true, the results should all be similar.

Klayman (11) has performed three-interval forced-choice experiments using some of the stimulus combinations described above. The results appear to exclude *proposition d*, since there is an indication of a difference between the click-cued light-detection and light-cued click-detection data. The averaged detectability-*vs.*- $\Delta t$  function obtained in these experiments displays a significant quadratic trend.

## REFERENCES

1. BOTHE, G. G., & MARKS, L. E. Absolute sensitivity to white noise under auxiliary visual stimulation. *Percept. & Psychophys.*, 1970, **8**, 176-178.
2. CHILD, I. L., & WENDT, G. R. The temporal course of the influence of visual stimulation upon the auditory threshold. *J. Exper. Psychol.*, 1938, **23**, 109-127.
3. EGAN, J. P., GREENBERG, G. Z., & SCHULMAN, A. I. Interval of time uncertainty in auditory detection. *J. Acoust. Soc. Amer.*, 1961, **33**, 771-778.
4. EGAN, J. P., SCHULMAN, A. I., & GREENBERG, G. Z. Memory for waveform and time uncertainty in auditory detection. *J. Acoust. Soc. Amer.*, 1961, **33**, 779-781.
5. GILBERT, G. M. Inter-sensory facilitation and inhibition. *J. Gen. Psychol.*, 1941, **24**, 381-407.
6. HENNEBERT, P. E. Influences des excitations lumineuses rythmées sur les fonctions de l'audition et de l'équilibre. *Acta-Oto-Rhino-Laryngol. Belg.*, 1955, **9**, 325-327.
7. HORN, G., & VENABLES, P. H. The effect of somaesthetic and acoustic stimuli on the threshold of fusion of paired light flashes in human subjects. *Quart. J. Exper. Psychol.*, 1964, **16**, 289-296.
8. HOWARTH, C. I., & TREISMAN, M. Lowering of an auditory threshold produced by a light signal occurring after the threshold stimulus. *Nature*, 1958, **182**, 1093-1094.
9. ———. The effect of warning interval on the electric phosphene and auditory thresholds. *Quart. J. Exper. Psychol.*, 1958, **10**, 130-141.
10. ———. Lowering of an auditory threshold by a near threshold warning signal. *Quart. J. Exper. Psychol.*, 1961, **13**, 12-18.
11. KLAYMAN, B. E. Detection and the timing of a heteromodal cue: Auditory-visual interaction. Unpublished Doctoral dissertation, Columbia University, New York, 1972.
12. KRAVCOV, S. V. The influence of sound upon the light and colour sensibility of the eye. *Acta Ophthal.*, 1936, **14**, 348-360.
13. KUROKI, S. The influence of light-stimulus upon hearing. *Jap. J. Psychol.*, 1937, **12**, 253-269. (English summary).
14. LONDON, I. D. Research on sensory interaction in the Soviet Union. *Psychol. Bull.*, 1954, **51**, 531-568.
15. MARUYAMA, K. The effect of intersensory tone stimulation on absolute light threshold. *Tohoku Psychol. Folia*, 1959, **17**, 51-81.
16. ———. "Contralateral relationship" between the ears and the halves of visual field in the sensory interaction. *Tohoku Psychol. Folia*, 1961, **19**, 81-92.
17. NEWHALL, S. M. Effects of attention on the intensity of cutaneous pressure and on visual brightness. *Arch. Psychol.*, 1923, **9**, Whole No. 61.
18. NOVAK, S. The effect of electrocutaneous digital stimulation on the detection of single and double flashes of light. (Doctoral dissertation, Columbia University) No. 63-1510, University Microfilms, Ann Arbor, Michigan, 1963.
19. ———. The effect of electrocutaneous digital stimulation on the detection of single and double flashes of light. *Psychol. Monog.*, 1965, **79**, Whole No. 608.
20. RYAN, T. A. Interrelations of the sensory system in perception. *Psychol. Bull.*, 1940, **37**, 659-698.
21. TREISMAN, M. The effect of one stimulus on the threshold for another: An application of signal detectability theory. *Brit. J. Stat. Psychol.*, 1964, **17**, 15-35.
22. TREISMAN, M., & HOWARTH, C. I. Changes in threshold level produced by a signal preceding or following the threshold stimulus. *Quart. J. Exper. Psychol.*, 1959, **11**, 129-142.
23. WATKINS, W. H. Effect of certain noises upon detection of visual signals. *J. Exper. Psychol.*, 1964, **67**, 72-75.

24. ———. Effect of intensity changes in auxiliary stimuli on auditory and visual signal detection. *USAF-ESD-TR-66-121*, United States Air Force, Bedford, Massachusetts, 1965.
25. ———. Photoc facilitation of tonal signal detection in a forced-choice situation. *Psychon. Sci.*, 1966, **6**, 477-478.
26. WATKINS, W. H., & FEEHRER, C. E. Investigations of acoustic effects upon visual signal detection. *USAF-ESD-TR-64-557*, United States Air Force, Bedford, Massachusetts, 1964.
27. ———. Acoustic facilitation of visual detection. *J. Exper. Psychol.*, 1965, **70**, 332-333.
28. WATKINS, W. H., & SCHJELDERUP, J. R. Effects of temporal variation of auxiliary light stimuli upon detectability of tonal signals. *Percept. & Psychophys.*, 1967, **2**, 317-322.
29. WATKINS, W. H., NICKERSON, R. S., & SCHJELDERUP, J. R. An automated facility for forced-choice signal detection experimentation. *USAF-ESD-TDR-64-383*, United States Air Force, Bedford, Massachusetts, 1964.

*New York State Psychiatric Institute*  
*722 West 168th Street*  
*New York, New York 10032*





## DETECTION AND THE TIMING OF A HETEROMODAL CUE: AUDITORY-VISUAL INTERACTION\*

*New York State Psychiatric Institute*

---

BRUCE E. KLAYMAN

---

### SUMMARY

In two three-interval temporal forced-choice experiments eight subjects were tested to determine the effect of a cue on the rate of detection of a signal when both cue and detection stimulus were of very brief duration and were separated by randomly varied temporal intervals (*delta ts*). *Delta ts* ranged from a cue precedence of 900 msec, through simultaneous presentation, to a detection-stimulus precedence of 900 msec.

A significant quadratic trend was found in the *delta-t* variable, with maximum detection when cue and detection stimulus were simultaneous, and there was an indication that the quadratic trend differed between the two stimulus combinations. An attempt was made to distinguish between attentional effects linked to the time of occurrence of the cue and attentional effects linked to the time of occurrence of the detection stimulus.

---

### A. INTRODUCTION

When a visual cue is varied in temporal position around an auditory detection stimulus, maximum sensitivity toward the detection stimulus seems to be achieved when the onset of the cue precedes the onset of the detection stimulus by between 0 and 500 msec; on either side of the maximum, sensitivity falls off, the decrease being more rapid in the direction of detection-stimulus precedence than in the direction of increasing cue precedence. There are not enough consistent data, however, to determine the shape of the function for the reciprocal condition—sound-cued light detection—or to compare maximum detection obtained under heteromodally cued conditions with detection in the absence of cue. [Experiments of the kind described above were reviewed in detail by Klayman (1).]

The present experiment attempts to compare light-cued-click-detection and click-cued-light-detection data obtained under comparable conditions, when

---

\*Recommended by Joseph Zubin of the Editorial Board, received in the Editorial Office on November 27, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

the cue is of very brief duration and varies from trial to trial around a similarly brief detection stimulus. A three-interval temporal forced-choice method was used, with three kinds of measures of no-cue "baseline" responding and two techniques for varying the difference between the times of detection-stimulus and cue onset (*delta t*).

## B. METHOD

### 1. Terminology and Conventions

In the present experiments the detection stimulus was presented between two electrocutaneous *markers* which signalled the subject to begin and to cease attending. However, in Experiment II the detection stimulus could occupy different temporal positions relative to the markers. Thus, the range of time that was occupied by the detection stimulus will be referred to as the *signal-occurrence interval* (SOI). The interval between the pair of markers that enclosed the SOI will be called the *marker-delimited interval* (MDI). The SOI had a fixed temporal relationship to each of the markers that defined the MDI.

In the present study the MDI was defined by markers 2400 msec apart, so that the duration of the MDI was 2400 msec. The detection stimulus was of very brief duration. In Experiment II its position within the MDI varied from 300 msec after the first electrocutaneous stimulus to 300 msec before the second. The SOI was thus about 1800 msec in duration. In the light-detection sessions of Experiment I, on the other hand, in which a 4-msec-duration detection stimulus was located always in the exact center of the MDI, the SOI was 4 msec long.

The timing of the cue will be expressed relative to the time of onset of the detection stimulus. By convention, the time of onset of the detection stimulus is taken as time zero, and when the onsets of the cue and of the detection stimulus were simultaneous, that situation will be referred to as zero time difference or  $\text{delta } t = 0$ . When the cue onset preceded the onset of the detection stimulus, the time difference *delta t* was *negative*; and when the onset of the cue followed the onset of the detection stimulus, *delta t* was *positive*. Trials in which cue was absent, when randomized together with cue-present trials, are referred to as plus-infinity trials ( $\text{delta } t = +\infty$ ), since in this case the onset of the cue may be regarded as indefinitely delayed.

### 2. Subjects

Eight Ss were employed: four men and four women ranging in age from 19 to 30 years, with a mean age of 23. There were two experiments and two

stimulus combinations: a man and a woman were tested in each of the four cells of this design.

The subjects were paid volunteers; all but one worked in the Columbia-Presbyterian Medical Center. Ss were pretested to ensure that they had normal auditory and visual sensitivity (wearing corrective lenses if necessary) and could focus on the stimulus patch.

### 3. *Apparatus*

The apparatus was designed to generate automatically stimulus presentations for three-interval temporal forced-choice experiments and to record stimulus conditions and S's response after each trial. Programming was effected between experiments and during intertrial intervals by means of manual switching of interconnections between various timers, pulse generators, or gating devices. The apparatus has been described in detail elsewhere (1).

### 4. *Procedure*

*a. Experimental setting.* The experiment was conducted in a ventilated, dimly lighted experimental room which contained an experimental booth. During testing the dark-adapted subject, wearing earphones, was seated in the booth, which was painted matte black inside, facing a hemispheric surround (integrating half-sphere) covered with diffusely reflecting white paint. The surround extended for  $180^\circ$  in the subject's visual field and contained a  $1^\circ 10'$  aperture—the stimulus patch—located in the midline at eye level. The uppermost portion of the hemisphere was directly illuminated, but the rest of it acted as a surround with a reasonably uniform luminance of .22 mlam (measured at the beginning of testing). The noise level in the booth was measured at 72 dB re SPL (ASA Scale C).

S sat in an adjustable chair with a backrest. His head was in an adjustable headrest with chin cup and forehead bar. The chin-cup contained two electrodes for delivering the electrocutaneous marker stimuli, mounted along the midline 18 millimeters apart. Each marker stimulus was a 4-msec pulse through S's chin of no more than 1-2 milliamperes. The current was maintained at an intensity that was clearly perceptible, but not painful for the individual subject.

S's right hand rested on a shelf, located below the hemisphere, on which there was a response-box containing a row of push-buttons. S pushed the button on the left to indicate the presence of a detection stimulus in the first MDI, the button in the middle to indicate its presence in the second MDI, and the button on the right to indicate the third MDI.

*b. General procedure.* The experiments consisted of temporal forced-choice (FC) trials which had the following form: Four electrocutaneous stimuli delimited three MDIs, each MDI 2400 msec in duration. The detection stimulus, which was present in every trial, was delivered in only one of the MDIs of a given trial. The cue, which was present in only "standard-block" trials was presented in all three MDIs of a given trial, in the same temporal position in each MDI; but the cue was randomly varied in temporal position from trial to trial.

The almost-instantaneous click cue was always set at 35 dB re SL, and the 4-msec-duration light cue was always at 25 dB re SL. The durations of the click- and light-detection stimuli were the same as above, but the intensity of the detection stimuli was always set at "threshold" (threshold being defined as the intensity that yielded about 2/3 correct detection in the absence of cues at the beginning of the experimental session) under the "initial-threshold" FC procedure.

Each subject was given one or two days of "preliminary-threshold" determination, during which his thresholds were obtained, by both the Forced Choice and Method of Limits techniques, to the standard click and light stimuli. Following this the subject received 12 experimental sessions, each session consisting of, in order of delivery, (a) "initial-threshold" trials—in each of which there was a detection stimulus but no cue—which were used to determine the detection-stimulus intensity to be employed during the remainder of the session; (b) two blocks of "standard" or "*delta-t*" trials, in which "finite-*delta-t*" trials, each containing a cue and a detection stimulus separated by a *delta t* which varied randomly from trial to trial, were randomized together with " $+\infty$ " trials, each of which contained a detection stimulus but no cue; (c) a block of "control" trials, each trial containing a detection stimulus but no cue; and (d) two more "standard" blocks. In each of the two experiments a single standard block consisted of a series of 24 trials in which each of the eight *delta ts* was presented randomly three times. A control block similarly contained 24 trials.

*Experiment 1.* In Experiment I the cue varied in temporal position within the MDI, while the temporal position of the detection stimulus was fixed. Specifically, the cue could occur in a given trial (if present at all) at either 300, 600, 1200, 1500, 1800, or 2100 msec after the electrocutaneous stimulus that marked the beginning of the MDI, while the detection stimulus only occurred 1200 msec after the beginning of the MDI. The *delta t* could therefore take the following values (in msec):  $-900$ ,  $-600$ ,  $-300$ ,  $0$ ,  $+300$ ,  $+600$ ,  $+900$ ,  $+\infty$ .



In this experiment two Ss were tested under the light-cued-FC-to-click condition and two under the click-cued-FC-to-light condition. Each S received a total of 432 initial-threshold trials (i.e., "criterion" trials, with the detection-stimulus intensity to be used in the remainder of the session), 288 control trials, and 144 trials under each of the eight *delta-t* conditions.

*Experiment II.* In Experiment II the detection stimulus varied in temporal position within the MDI, while the temporal position of the cue was fixed. Specifically, in a given trial the detection stimulus could occur at either 300, 600, 900, 1200, 1500, 1800, or 2100 msec after the beginning of the MDI; while the cue only occurred 1200 msec after the MDI beginning, when it occurred at all. The *delta t* could therefore take the following values (in msec):  $+\infty$ , +900, +600, +300, 0, -300, -600, -900. Thus—whereas in Experiment I a *delta t* of -900 meant that the cue occurred 300 msec after the beginning of the MDI, and the detection stimulus occurred 1200 msec after the beginning of the MDI—in Experiment II the same *delta t* signified that the cue occurred 1200 msec after the beginning of the MDI, and the detection stimulus occurred 2100 msec after the MDI beginning. (In the case of the  $+\infty$  trials, in which cue was absent, the detection stimulus randomly and equally often took the seven temporal positions indicated above.)

In Experiment II two Ss were tested under the light-cued-FC-to-click condition and two under the click-cued-FC-to-light condition. In this experiment each S received a total of 504 initial-threshold criterion trials, 288 control trials, and 144 trials under each of the eight *delta-t* conditions.

## C. RESULTS

### 1. Analysis of the Data

The experimental data—all 12 sessions for each of the eight subjects—were submitted to two Analyses of Variance and an Analysis of Quadratic Trend. The data employed in each Analysis were *arcsin-transformed proportions*, each proportion being obtained by dividing the number of correct detections of the critical MDI within a given condition by the total number of trials within that condition. The model for ANOVAR I and II is presented in Winer (2, p. 337 ff.), and the model for the Quadratic Trend Analysis also (p. 353 ff.).

*a. ANOVAR I.* In this analysis (presented in Table 1) the following main effects were investigated, along with appropriate error terms and interactions: "Modalities"—light-cued click detection *vs.* click-cued light detection; "Experiments"—Experiment I (cue position varying) *vs.* Experiment II (detection-stimulus position varying); and "*Delta ts*"—*delta t* = -900, -600, -300, 0, +300, +600, or +900 msec (*delta t* =  $+\infty$ , a cue-absent condi-



TABLE 1  
SUMMARIES OF ANALYSES OF VARIANCE AND ANALYSIS OF QUADRATIC TREND

Source of variation	SS	df	MS	F
<i>ANOVAR I</i>				
Between subjects	.7211	7		
Modalities (A)	.0011	1	.0011	.0120
Experiments (B)	.0307	1	.0307	.3348
AB	.3225	1	.3225	3.5169
Error between	.3668	4	.0917	
Within subjects	.7099	48		
<i>Delta ts</i> (C)	.1871	6	.0312	2.2286*
AC	.0858	6	.0143	1.0214
BC	.0426	6	.0071	.5071
ABC	.0581	6	.0097	.6929
C $\times$ Error within	.3363	24	.0140	
<i>Analysis of Quadratic Trend</i>				
Within subjects (quadratic)	.1477	8		
C (quadratic)	.1232	1	.1232	68.4444**
AC (quadratic)	.0132	1	.0132	7.3333*
BC (quadratic)	.0035	1	.0035	1.9444
ABC (quadratic)	.0006	1	.0006	.3333
C $\times$ Error within (quadratic)	.0072	4	.0018	
<i>ANOVAR II</i>				
Between subjects	.5014	7		
Modalities (A)	.0116	1	.0116	.1734
Experiments (B)	.1288	1	.1288	1.9253
AB	.0935	1	.0935	1.3976
Error between	.2675	4	.0669	
Within subjects	.1700	24		
Conditions (C)	.0318	3	.0106	1.9273
AC	.0078	3	.0026	.4727
BC	.0266	3	.0089	1.6182
ABC	.0379	3	.0126	2.2909
C $\times$ Error within	.0659	12	.0055	

Note: Sources of variation are defined as follows: Modalities (A)—i.e., light-cued click detection or click-cued light detection. Experiments (B)—i.e., Experiment I (cue position varying) vs. Experiment II (detection-stimulus position varying). In ANOVAR I and the Analysis of Quadratic Trend, "C" is defined as *Delta ts*—i.e., the seven cue-present conditions (—900, —600, —300, 0, +300, +600, and +900 msec). In ANOVAR II, "C" is defined as Conditions—i.e., the "cue-present" condition (the seven *delta-t* conditions of the preceding analyses combined) plus the three "cue-absent" conditions: "initial threshold" (no-cue criterion trials presented in blocks at the beginning of the session), "control" (no-cue trials presented in blocks in the middle of the session), and "+ $\infty$ " (no-cue trials randomized together with the cue-present trials).

\*  $p < .10$ .

\*\*  $p < .01$ .

tion, was not included). The results of this analysis show that the *delta-t* variable approaches significance at the .05 level, while no other source of variation approaches significance.

b. *Quadratic Trend Analysis.* The *delta-t* variable which was found to be

nearly significant in ANOVAR I was analyzed further: Figure 1, which depicts the data analyzed in ANOVAR I, presents detectability as a function of  $\Delta t$  (with the  $+\infty$  point omitted, as an ANOVAR I). In this figure, and in the figure that follows,  $\Delta t$  is depicted on the abscissa, while the ordinate variable is "Percent correct". "Percent correct" is defined as 100 times the number of correct detections of the critical MDI divided by the total number of trials in each  $\Delta t$ -condition; it was obtained, without correction for chance guessing by collapsing over the three MDIs. Figure 1

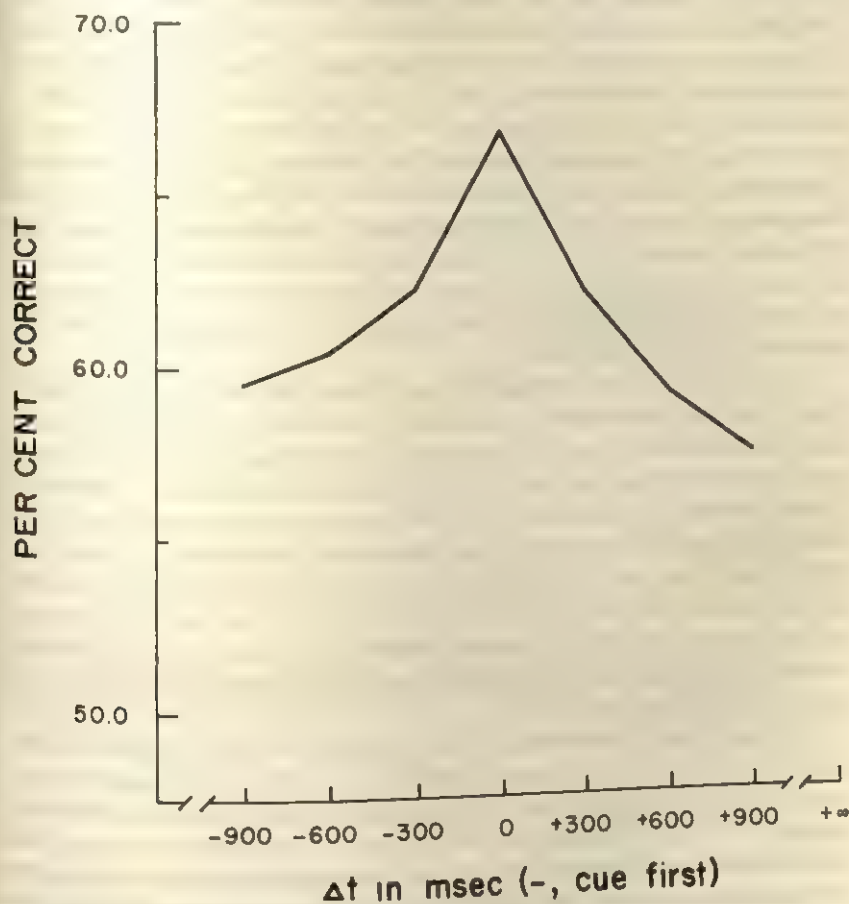


FIGURE 1  
OVERALL PERCENT CORRECT AS A FUNCTION OF  $\Delta t$  FOR THE EIGHT SUBJECTS  
 $N = 1152$  for each  $\Delta t$ . (The  $+\infty$  condition has been omitted.)

depicts an upward-convex, curvilinear function, peaking at  $\text{delta } t = 0$ , which is rather symmetrical and without reversals. This function seems describable by a quadratic equation, and for this reason an Analysis of Quadratic Trend was performed on the data included in ANOVAR I.

In the Quadratic Trend Analysis (presented in Table 1) the main effect ("*Delta ts*"), error term, and interactions are the same as in the "Within subjects" part of ANOVAR I (Table 1). The results indicate a highly significant quadratic trend in the *delta-t* variable ( $p < .01$ ), demonstrating that detectability (percent correct) as a function of *delta t* between  $-900$  and  $+900$  msec is describable by a quadratic equation. Furthermore, the modalities-by-*delta-ts* interaction approaches significance at the .05 level. This interaction suggests a real difference, since a "relaxed" criterion would seem to be advisable in view of the small number ( $N = 2$ ) of subjects in each cell and the small numbers of degrees of freedom. The modalities-by-*delta-ts* interaction is depicted in Figure 2, in the form of averaged detectability-vs.-*delta-t* functions presented separately for the light-cued-FC-to-click and the click-cued-FC-to-light subjects. It may be noted that the quadratic trend seems to be more marked in the case of the light-cued-click-detection curve than in the case of the click-cued-light-detection curve, which latter appears relatively flat between  $-900$  and  $+300$  msec.

c. ANOVAR II. In this analysis the following main effects were investigated, along with appropriate error terms and interactions (Table 1): "Modalities," "Experiments," and "Conditions." "Modalities" and "Experiments" were defined in connection with ANOVAR I, and there were four "Conditions"—the "cue-present" condition (the combined seven *delta-t* conditions of the preceding Analyses) plus the three "cue-absent" conditions. The three cue-absent conditions were "initial-threshold" (no-cue criterion trials presented in blocks at the beginning of the session), "control" (no-cue trials presented in blocks in the middle of the session), and " $+\infty$ " (no-cue trials randomized together with the cue-present trials). None of the sources of variation in this analysis approaches significance.

The result that "Conditions" is not a significant source of variation is taken as an indication that performance did not vary significantly among the three cue-absent conditions and the cue-present condition. However, the cue-present condition in ANOVAR II was derived by pooling across *delta ts*, thus eliminating *delta t* itself as a statistical variable. Therefore, the finding that average performance across *delta ts* in the pooled cue-present condition does not differ from performance in the cue-absent conditions does not contradict the effect of *delta t* in the cue-present condition.

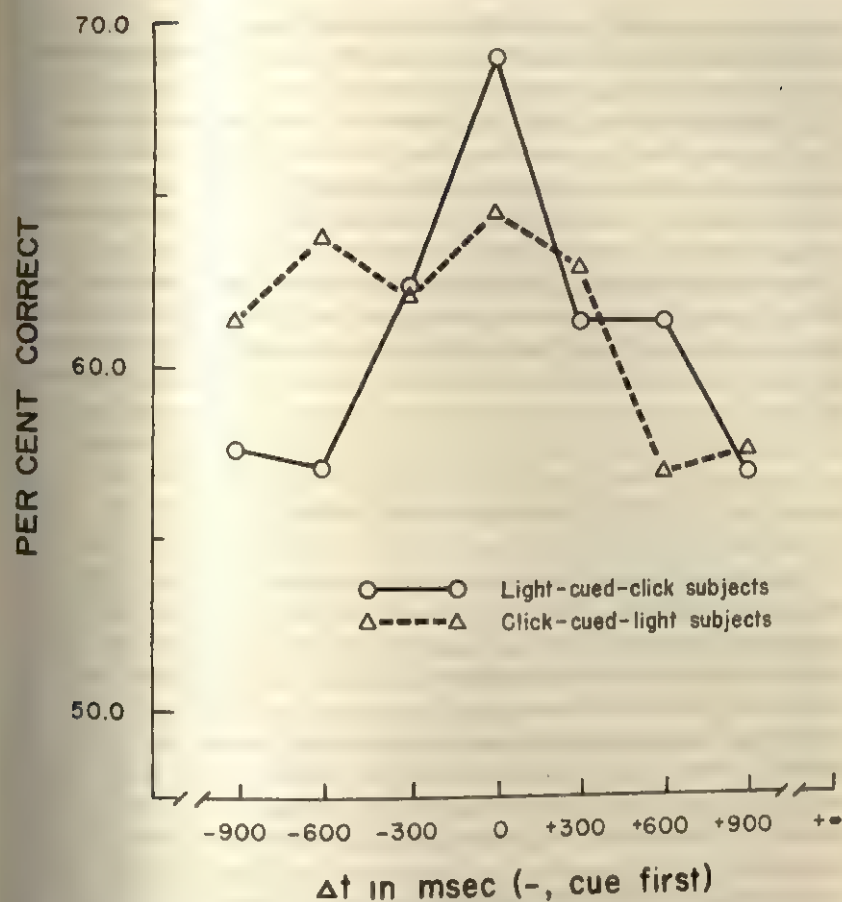


FIGURE 2

PERCENT CORRECT AS A FUNCTION OF  $\Delta t$  PRESENTED SEPARATELY FOR THE LIGHT-CUED-CLICK-DETECTION AND CLICK-CUED-LIGHT-DETECTION SUBJECTS  
 $N = 576$  for each  $\Delta t$  point.

## 2. Conclusions

a. *Quadratic trend in delta-t variable.* A significant quadratic trend was found in the delta- $t$  variable; that is, detection rate as a function of  $\Delta t$  between  $-900$  and  $+900$  msec was found to be describable by a quadratic equation. This quadratic trend, depicted in Figure 1, seems to take the form of a continuous rise in detection rate as the interval of time by which the heteromodal cue precedes the detection stimulus is reduced from 900 msec to

0 (simultaneity), and of a (slightly faster) continuous decline in detection rate as the interval of time by which the heteromodal cue follows the detection stimulus increases from 0 msec to 900 msec.

*b. Other indications.* There is a suggestion that the quadratic equation which best describes the *delta-t* variable differs between the light-cued-click-detection and the click-cued-light-detection data. There is no indication of any difference between results obtained in Experiment I (cue position varying) and Experiment II (detection-stimulus position varying).

There is no indication of any difference among the three "cue absent" conditions (initial threshold, control, and  $+\infty$ ) or between the various cue-absent conditions and the "cue-present" data as a whole. Thus it cannot be said that the presence of a heteromodal cue was in itself either inhibitory or facilitory.

#### D. DISCUSSION

##### 1. *What Is the Effective Independent Variable?*

The lack of detectable difference between the results of Experiments I and II—in none of the three analyses did the "Experiments" variable, or any of its interactions, even approach significance—suggests the nature of the independent variable employed in the present study. In Experiment I the temporal position of the cue was varied, while that of the detection stimulus remained fixed; in Experiment II the temporal position of the detection stimulus was varied, while that of the cue remained fixed. Thus neither variation of cue position nor variation of detection-stimulus position was in itself a necessary condition of the experimental results obtained. Furthermore, each *delta t* in Experiment II was a mirror-image of the corresponding *delta t* in Experiment I, in the sense that a *delta t* which occurred in the first half of the 2400-msec-duration MDI in one experiment occurred in the second half of the MDI in the other experiment (see *Procedure: Experiment I*). Thus the location of the cue-detection stimulus pair within the MDI does not seem to be a relevant variable in determining detection rate under the present conditions. It appears to be the temporal relationship between the members of the cue-detection stimulus pair (the *delta t*) which is the effective independent variable.

##### 2. *Time-Specification vs. Arousal and Distraction*

In a review of the literature on heteromodally cued auditory and visual detection experiments (1) it was noted that experimenters who employed a brief or momentary heteromodal cue tended to favor one of two types of



"attentional" explanation of the effects of the cue: (a) specification of the time of detection-stimulus occurrence, or (b) arousal effects of the cue. It was impossible to decide, on the basis of the results of previous experiments, which of the two alternative types of attentional explanation—time-specification or arousal—was more important in the case of a momentary cue varying from trial to trial in temporal position around a detection stimulus. This problem can be more broadly restated as follows: Which are more important in determining the shape of the detectability-*vs.*- $\delta t$  function—processes that lead to heightened or reduced attentiveness to the arrival of the detection stimulus (a) near the expected time of occurrence of the detection stimulus or (b) near the time of occurrence of the cue? The kind of explanation that is based on heightened or reduced attentiveness near the expected time of detection-stimulus occurrence will be called the *time-specification hypothesis*. The kind of explanation that is based on effects of the cue may involve either heightened attentiveness near the time of occurrence of the cue (*arousal hypothesis*) or reduced attentiveness near the time of occurrence of the cue (*distraction hypothesis*).

Among the alternative attentional explanations the data do not favor an explanation based exclusively on the time-specification hypothesis. The reason for this conclusion is the lack of difference between the results of Experiments I and II. Since in Experiment I the detection stimulus (4 msec or less in duration) was always presented in the exact center of the MDI, while in Experiment II the temporal position of the detection stimulus varied from trial to trial within the 1800-msec SOI, the time of detection-stimulus occurrence was much better specified in Experiment I than in Experiment II. If the time-specification hypothesis applied exclusively, one would expect a difference between the two sets of data.

a. *Is cue-first data determined by arousal or distraction?* A plausible version of the arousal and distraction hypotheses would state that, in trials where the cue precedes the detection stimulus, shortly after the onset of the cue the level of attentiveness gradually increases (arousal) or decreases (distraction) to a point or plateau of maximal or minimal attentiveness, then gradually returns to a "resting" level. This process would be reflected in a rise in detection rate to a maximum, followed by a decline, or in a decline in detection rate to a minimum, followed by a rise, as  $\delta t$  varies from some large negative value to zero—i.e., from considerable cue precedence to simultaneous presentation of cue and detection stimulus. In either the arousal or distraction case the rise and decline need not both require the same amount of time.

In the negative- $\delta t$  segment of Figure 1 (i.e., that part of the curve

extending from  $-900$  msec to  $0$ ) detectability increases at an accelerating rate as  $\delta t$  varies from maximal cue precedence (i.e.,  $-900$  msec) to simultaneity ( $0$  msec). This finding may be explained in two ways: (a) by *arousal*—an increase in detection rate to a maximum located between  $-300$  msec and  $0$ , and a decrease between the maximum and  $0$ ; or (b) by *distraction*—starting with a  $\delta t$  greater than  $-900$  msec, a decrease from that point to a minimum near  $-900$ , and an increase between the minimum and  $0$ .

To be able to choose with any degree of certainty between the arousal and distraction alternatives, an experiment would be required in which  $\delta t$ s more negative than  $-900$  msec, and  $\delta t$ s between  $-300$  and  $0$ , would be added to the set of  $\delta t$ -conditions here employed. It is of course possible that both explanations are valid—for some of the eight individual-subject curves which were averaged to produce Figure 1 the negative- $\delta t$  segments may be determined by arousal, for others they may be determined by distraction; alternatively, the negative- $\delta t$  segments of some or all of the individual-subject curves may be determined by a complex mixture of both processes.

b. *Is cue-second data determined by time-specification?* An arousal or distraction hypothesis, as defined above, seems to be a less plausible explanation of the retroactive effect of the cue than would be a time-specification hypothesis. Such a hypothesis would explain in slightly different ways the contributions of Experiment I and Experiment II to the positive- $\delta t$  segment of the overall function Figure 1, which displays a continuous decline in detection rate between  $\delta t = 0$  and  $\delta t = +900$  msec.

In *Experiment II* the temporal position of the detection stimulus within the MDI was varied from trial to trial, while the cue was fixed at the center of the MDI. In this experiment attentiveness to the arrival of the detection stimulus may have risen continuously from the beginning to the center of the MDI, which latter was always marked by the cue, when cue was present. The decrease in detection rate as a function of increasing detection-stimulus precedence would thus reflect in a direct way the change in attentiveness in the first half of the MDI: The data point at  $\delta t = 0$  would be determined by attentiveness at the center of the MDI (the point in time at which the detection stimulus occurs in the  $\delta t = 0$  condition). Similarly, the data point at  $\delta t = +300$  msec would be determined by the lower level of attentiveness 300 msec before the center of the MDI; the data point at  $+600$  msec by the still lower attentiveness 600 msec before the MDI center; etc.

In *Experiment I* the position of the cue was varied, while the cue was

fixed at the MDI center. In this experiment the center of the MDI was not marked by the cue on most trials and thus could not have been estimated with complete accuracy; it may have been re-estimated in every trial. It is possible that the point in time chosen as the estimate of the center of the MDI was affected by the presence of a cue in the second half of the MDI, perhaps because in the first MDI of a given trial the presence of a cue in the second half was noted, and thus influenced the estimate of the center of the second and third MDIs. The influence that a cue in the second half of an MDI may exert on the estimate of the center of a subsequent MDI in the same trial cannot be postulated with any degree of assurance, but one can visualize a time-curve of attentiveness which peaks at the estimated center of the MDI and moves about the temporally fixed detection stimulus, as  $\delta t$  is varied, in such a way as to produce a lowering of detection rate with increasingly large positive  $\delta t$ s.

It should be noted that in the first half of each MDI in Experiment I and the second half of each MDI in Experiment II (in both cases the location of cue-first [negative]  $\delta t$ s) changes in attentiveness due to time-specification processes may interact complexly with changes due to arousal or distraction.

#### REFERENCES

1. KLAYMAN, B. E. Detection and the timing of a heteromodal cue: Auditory-visual interaction. Unpublished Doctoral dissertation, Columbia University, New York, 1972.
2. WINER, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.

*New York State Psychiatric Institute*  
*722 West 168th Street*  
*New York, New York 10032*



## FACTOR SIMILARITY OF PERSONALITY ACROSS PRIVATE AND MILITARY SAMPLES: AN ANALYSIS OF THE PERSONALITY/ATTITUDE SCHEDULE\*<sup>1</sup>

*Graduate School of Business and Industrial Relations Research Institute,  
The University of Wisconsin-Madison; and Graduate School  
of Business, Indiana University*

L. L. CUMMINGS, DONALD L. HARNETT,<sup>2</sup> AND STUART M. SCHMIDT

### SUMMARY

The use of personality and attitudinal questionnaires in private executive and military samples is briefly reviewed. The need for evidence regarding the factor similarity of such questionnaires in comparative studies is noted. Such a comparative factorial study of an abbreviated version of the Shure-Meeker (11) Personality/Attitude Schedule (PAS) is reported. In terms of the general scales examined, three of the four scales emerge in both samples. The order of emergence, as well as the number and item content of the factors, differs slightly between the two samples. Caution is suggested in assuming factor similarity across samples in comparative studies utilizing the PAS and other personality and attitudinal instruments.

### A. INTRODUCTION

This study seeks to determine the factor structure similarity of certain personality dimensions of comparable groups of private sector business executives and military officers. There has been previous research comparing the perceived importance and satisfaction of needs (as well as perceived role requirements) in business and military samples (9, 10). These studies have found that military officers perceived greater need dissatisfaction than their civilian counterparts and that the military officers shared with the civilian executives the perception that inner-directed personality traits are more important for job success than other-directed traits.

In a different context and for different purposes, Vinacke (13) reviewed

\* Received in the Editorial Office on November 28, 1972, and published at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The financial support of the Center for International Business Research, Graduate School of Business, University of Wisconsin-Madison, and the Ford Foundation (through the International Business Research Institute, Graduate School of Business, Indiana University) is gratefully acknowledged.

<sup>2</sup> The second author is at Indiana University.



several experimental studies examining the effects of task, situational, and personality variables on bargaining behavior. These studies have exhibited the impact of several general categories of personality variables on bargaining; namely, psychopathology, attitudes and traits, and motives. In particular, the following personality characteristics have been found to be associated with behavior in bargaining situations: authoritarianism (12), dogmatism (3), propensity to trust (14), levels of ascendance (4), level of aspiration (7), machiavellianism (1), and risk taking (6).

One important missing element in both of these streams of research is a systematic effort to assess the comparability of the measuring instruments used across samples. This is true in both the comparative need satisfaction and personality studies. The essential question is whether a measuring instrument is assessing similar constructs or variables when comparing personality and satisfaction across samples.

Cummings, Harnett, and Schmidt (2) have reviewed the evidence on this problem in the context of cross-cultural personality assessment and have found moderate support for the cross-cultural generality of the Shure-Meeker (11) Personality/Attitude Schedule (PAS). The present study examines the factor stability of the PAS across the military-private sector of managerial or executive employment. Our concern here is the extent to which professional executives in the military (career officers) and in the private industrial employment sector exhibit different factor structures of personality.

The PAS was designed to assess six basic dimensions of personality thought to be particularly relevant determinants of interpersonal bargaining behavior. Our interest, both in this study and in the managerial cross-cultural study (2), focuses upon four of these dimensions: (a) tendency toward belligerence *versus* conciliation in interpersonal relations, (b) tendency toward risk taking *versus* risk averting behavior, (c) belief in control by fate *versus* self or internal control, and (d) tendency toward suspiciousness *versus* trust in interpersonal relations.

## B. METHOD

### 1. Instrument

In the development of the PAS, Shure and Meeker (11) selected, from commonly used test instruments, 24 scales on the basis of their relevance to bargaining behavior. From these scales, six new orthogonal factors evolved through the use of factor analysis of the original tests. Then for all the subjects, factor scores were obtained which were used in obtaining validity coefficients for all the test items. In this manner, sets of items with high

validities on only one of the six factor score criteria were identified. The final selection of items in the construction of the factor scales started with the elimination from the item pool of those items with multiple high validities on more than one criterion.

Following this stage of development, 129 of the most promising items were identified by the Wherry-Gaylord method of selecting items, making the validity coefficient between the items with fixed weights and the factor score criteria as high as possible. Thus Shure and Meeker (11) are presumed to have obtained the items with the highest validity for each scale which add cumulatively to the prediction of each criterion.

The next step in their test development was to factor analyze the 129 item scores obtained from 544 freshman and sophomore college students. Of the 129 items, 27 had low loadings on all the scales and were consequently eliminated. The remaining 102 items were assigned to the six scales which were designated as Aggressive militarism, Conciliation *versus* belligerence in interpersonal relations, Authoritarian nationalism *versus* equalitarian internationalism, Risk avoidance *versus* risk taking, External *versus* internal control, and Suspiciousness *versus* trust.

For this study, four of these six scales were chosen on the basis of relatedness to interpersonal bargaining behavior. Thus a 58 item questionnaire was adapted from the original Shure and Meeker (11) Personality/Attitude Schedule. This instrument includes the dimensions of Conciliation *versus* belligerence in interpersonal relations, Risk avoidance *versus* risk taking, External *versus* internal control, and Suspiciousness *versus* trust.<sup>3</sup> These four dimensions were measured by means of seven-point and three-point scales, as well as through forced-choice questions. Each of the four dimensions along with an illustrative question is described below.

a. *Conciliation versus belligerence in interpersonal relations.* "High scorers advocate responding to the needy and less fortunate (or even to unfriendly, quarrelsome, provocative, and hostile persons) with understanding, help, and friendliness. They urge admitting their own wrongs, and refuse to use threats or belligerent means or to be motivated by revenge. Instead, they advise a diplomatic and constructive response guided by considerations of humanitarianism and cooperation" (11, p. 240). Low scorers exhibit the reverse tendencies and are characterized as belligerent. An illustrative question for this dimension is: "When someone has been nasty to you, you should try to understand what's bothering him, so that you can be helpful."

<sup>3</sup> Copies of the modified PAS questionnaire can be obtained from the first author at the address shown at the end of this article.

On this type of question the executive was instructed to indicate, on a seven-point scale, the degree to which he agrees with this statement (from "I agree very much" to "I disagree very much"). Total scores on this scale can range from 79 (highest conciliation) to -23 (highest belligerence) with a midpoint of 28.

*b. Risk avoidance versus risk taking.* "High scorers are unadventuresome, exhibit a low activity level, and are unwilling to expose themselves to dangers or to hazard risks of either a material or physical character" (11, p. 241). Low scorers can be characterized in a reverse fashion and might be termed risk takers. An illustrative question is: "Do you like to invest money in a promising invention?" For this type of question the executive was asked to answer either *yes* (a score of 1), *cannot decide* (a score of 2), or *no* (a score of 3). The potential extremes on this scale are 48 (highest risk avoidance) to 16 (highest risk taking) with a midpoint of 32.

*c. External versus internal control.* "High scorers believe that events are controlled by external forces (fate, chance, events). Low scorers believe that they can exercise some control over the events around them" (11, p. 241). An illustrative question is the following: "(a) People's misfortunes usually result from the mistakes they make or (b) sometimes I feel that I don't have enough control over what happens to me." The executive was asked to indicate which of the two statements he more strongly believed to be the case. The possible extremes on this scale are from +6 (highest belief in external control) to -3 (highest belief in internal control) with a theoretical midpoint of +1.5.

*d. Suspiciousness versus trust.* "High scorers are characterized by . . . selfishness, projection of hostility, excitability, and tenseness. Low scorers are characterized by a trusting, unselfish, calm, and optimistic orientation" (11, p. 241). An illustrative question is: "Even people who appear friendly to you may be unreliable because they are mainly concerned with their own interests." The executive was asked to respond on a seven-point scale, ranging from "I agree very much" (a score of 7) to "I disagree very much" (a score of 1). Possible total scores on this scale can range from 76 (highest suspiciousness) to -12 (highest trust) with a midpoint of 32.

## 2. Subjects

The adapted PAS was completed by 51 male executives in advanced management courses at Indiana University and Harvard University. The military officer group was composed of 66 majors and colonels attending courses at the Industrial College of the Armed Forces. These military officers held com-

parable administrative positions, in terms of organizational rank, to their private sector executive counterparts.

The relative equivalence of the military and private groups should be noted. The average age of the military officers was 44.5 years, and that of the private executives was 44.8 years. The average number of years experience in the military was 22.3 years, and the average number of years experience in business for the private executives was 20.1 years. The military officers averaged 5.5 years of formal education beyond age 18, while the private executives averaged 3.9 years.

### C. RESULTS

The military and private executive groups were factor analyzed separately with the results presented in Table 1. The factor analytic procedure used for the 58 item PAS was equivalent to that used in the Shure and Meeker (11) PAS cross-validation studies. This method consisted of (a) constructing a correlation matrix for all the item scores, (b) factoring with use of the principal axis solution, and (c) rotating these factors with use of a varimax procedure (8).<sup>4</sup> For each factor that emerged in our analysis, only loadings of .40 or above are reported. The PAS items are designated by number in reporting the results of the analysis.<sup>5</sup> Seven primary factors account for 44.3 percent of the total variance and 56.2 percent of the total factors variance in the military group, and 43.9 percent and 52.8 percent, respectively, in the private executive group.

Table 2 presents a summary of the interpretations of the seven factors which, after extraction and rotation, were interpretable for either one or both samples.

#### 1. Interpretation of Factors

The first factor extracted is interpreted as a conciliation *versus* belligerence factor in both the military and private executive groups and is principally defined by loadings of the PAS conciliation-belligerence items and on several suspiciousness-trust items (items 25, 8, and 12 for the military, and 16, 25, and 12 for the private executives).

The second factor is interpreted differently for the military officers and the private executives. For the military group, this factor is interpreted as being subject to fate in a passive manner or taking chances in a passive fate sense

<sup>4</sup> A STATJOB Factor 2 program was used on a UNIVAC 1108 computer. The assistance of the staff of the Social Systems Research Institute, University of Wisconsin, is gratefully acknowledged.

<sup>5</sup> The item numbers refer to designations in the modified PAS questionnaire.



TABLE 1  
ITEM LOADINGS AND RANK BY FACTOR AND GROUP (MILITARY AND PRIVATE EXECUTIVE)

Rank	I				II				III			
	Milit		Priv		Milit		Priv		Milit		Priv	
	It	Lo	It	Lo	It	Lo	It	Lo	It	Lo	It	Lo
1	15	875	27	878	44	783	31	837	31	802	3	829
2	7	855	29	876	56	—776	37	775	41	783	9	816
3	29	815	15	868	55	—714	41	728	37	673	18	627
4	27	798	7	837	58	653	42	660	32	654	11	580
5	1	759	13	801	45	580	34	603	34	623	19	481
6	24	747	17	799	43	—568			42	576		
7	28	746	24	786								
8	17	728	28	755								
9	25	—670	22	686								
10	20	598	1	669								
11	5	579	5	634								
12	22	559	16	—606								
13	8	503	25	—511								
14	11	492	18	479								
15	13	487	11	468								
16	12	—481	12	—403								
% total factor variance <sup>a</sup>												
	19.0		19.7		8.9		6.9		7.8		6.3	
% total variance <sup>b</sup>												
	15.0		16.4		7.0		5.7		6.2		5.2	

Note: It = item; Lo = loading.

<sup>a</sup> Sum of total factor variance for military sample = 56.2%; for private sample = 52.8%.

<sup>b</sup> Sum of total variance for military sample = 44.3%; for private sample = 43.9%.

and is defined by loadings of the PAS risk-taking items 44 and 45, the internal-external control items 56, 55, and 58, and the suspiciousness-trust item 43. For the private executives, Factor II is interpreted as a risk-taking factor also, but it refers to the taking of risks involving an element of personal danger. The factor is defined by loadings of the PAS risk-taking items 31, 37, 41, 42, and 34.

The third factor extracted is also interpreted differently between the two groups. For the military group, this factor refers to the taking of risks involving an element of personal danger. It is essentially the same factor as Factor II for the private executive group and is defined by loadings of the PAS items 31, 41, 37, 32, 34, and 42. The third factor for the private executives is interpreted as conciliation-belligerence in relation to competitive and helping situations and is defined by loadings of the PAS conciliation-belligerence items 3, 9, 18, and 11, and the suspiciousness-trust item 19.

The fourth factor for the military group is interpreted as a risk-taking factor but specific to risks involving monetary stakes. This factor is defined



TABLE 1 (continued)

IV				V				VI				VII			
Milit		Priv		Milit		Priv		Milit		Priv		Milit		Priv	
It	Lo	It	Lo	It	Lo	It	Lo	It	Lo	It	Lo	It	Lo	It	Lo
30	839	58	858	36	634	47	874	14	741	35	756	47	819	20	722
38	714	39	—765	37	425	2	577	13	685	32	715	35	656	56	—670
33	702					14	487	16	—474	55	642	39	553	26	—605
						8	448							8	419
						6	432								

5.6                  5.1                  5.4                  5.0                  4.8                  4.9                  4.7                  4.9

4.4                  4.3                  4.2                  4.1                  3.8                  4.1                  3.7                  4.1

by loadings of the PAS risk-taking items 30, 38, and 33. For the private executives, Factor IV is not interpretable and only consists of loadings of the PAS internal-external control item 58 and the suspiciousness-trust item 39.

The fifth factor differentiated between the military officers and private executive groups. For the military officers, this factor is interpreted as taking risks involving speed with a vehicle and is defined by loadings of the PAS items 36 and 37. The fifth factor for the private executives is interpreted as suspiciousness-trust in interpersonal judgments and is defined by loadings of the items 47, 2, 14, 8, and 6.

The sixth factor for the military officer group is interpreted as mercenary, hostile interpersonal relations and is defined by loadings of the PAS suspiciousness-trust items 14 and 16 and the conciliation-belligerence item 13. For the private executives, this factor is not interpretable and consists of loadings of the PAS suspiciousness-trust item 35, risk-taking item 32, and internal-external control item 55.

The seventh factor is interpreted for the military officers as a suspicious-

TABLE 2  
INTERPRETATIONS OF THE FACTOR COMPOSITION BASED ON ITEM CONTENT

Factor	Military sample	Private executive sample
I	Conciliation-belligerence	Conciliation-belligerence
II	Refers to being subject to fate in a passive manner; taking chances in a passive fate sense	Taking risks involving an element of personal danger
III	Taking risks involving an element of personal danger	Conciliation-belligerence in relation to competition and helping
IV	Taking risks involving money; "gambling"	Not interpretable
V	Taking risks involving a car, speed: "Dragster Factor"	Suspiciousness-trust in interpersonal judgments
VI	Mercenary, hostility in interpersonal relations	Not Interpretable
VII	Suspiciousness-trust relating to personal emotions; "moods"	Not Interpretable

ness-trust dimension relating to personal emotions or moods. This factor is defined by loadings of the PAS suspiciousness-trust items 47, 35, and 39. The seventh factor is not interpretable for the private executive group and consists of loadings of the PAS conciliation-belligerence items 20 and 26, internal-external control item 56, and the suspiciousness-trust item 8.

From this factor analysis of the military officer group and the private sector executive group, we see that the first factor is virtually identical in both groups. The conciliation-belligerence dimension accounts for a large proportion of the variance in these two groups. However, on the other factors, the military officers appear to manifest a different structural profile from that of the private executives. The military structure consists of dimensions of passive subjection to fate, three forms of risk taking, hostility in interpersonal relations, and moods of suspiciousness-trust. For the military officer group, risk avoidance *versus* risk taking is differentiated into risks involving personal danger, monetary stakes, and vehicular speed. The private executives' structure contains only a risk-avoidance *versus* risk-taking factor involving an element of personal danger. The other two factors that are interpreted for the private executives are conciliation *versus* belligerence in relation to competition and helping situations, and the factor of suspiciousness *versus* trust in interpersonal judgments.

What can we conclude concerning the factor similarity of personality in the two samples as assessed by the PAS? Two specific questions can be asked in this regard. First, which of the original four PAS scales as extracted by

Shure and Meeker (11) emerged in these samples? Second, what specific differences are evident between the two samples?

The *conciliation-belligerence* scale emerged in both samples and in both cases emerged first after rotation, thus accounting for the greatest amount of factor and total variance. This scale and the underlying personality construct appear to be similar and stable across the two samples. The *suspiciousness-trust* scale emerged in both samples but with slightly different item compositions. In the military sample the scale seems to be specific to mercenary hostility and to a general emotional tone or "mood" in interpersonal relations. In the private executive sample, the scale emerges in a less specific form with an item composition similar to the original Shure-Meeker PAS scale. The *risk taking-risk avoidance* scale emerges in both samples. However, this personality characteristic is more differentiated in the military than in the private sample. Risk taking in the private sample is focused, as an identifiable factor, on physical danger involving risk to physical well being. The *internal-external control* scale emerges only in the military sample. It does not emerge as an identifiable characteristic of the private executive sample.

In general, more interpretable factors are identifiable for the military officers than the private executives. In addition, there exists greater sub-factoring or differentiation within the military sample.

## 2. Congruence Among Factors

In the comparison of the factor structures obtained from the military and private executive samples, it is of interest to note the degree of statistical congruence between the respective factors. Since the same variables are used in both factor analyses, it is possible to determine the degree of agreement between like factors in each sample. This is accomplished by use of the factor weights to determine the coefficient of congruence between any two factors (5, p. 270). This index takes the summations of the factor weights over the number of variables, given different samples, to yield a measure of agreement between factors. The coefficient of congruence can range from +1 in the case of perfect agreement, to zero when there is no agreement, to -1 in the case of perfect inverse agreement.

In this study we computed the coefficients of congruence for the first seven factors of each sample, respectively. As shown in Table 3 each of the seven factors from the military sample is compared with all of the factors from the private executive sample.

TABLE 3  
COEFFICIENT OF CONGRUENCE MATRIX

Factors I to VII from private executive sample	Factors I to VII from military sample						
	I	II	III	IV	V	VI	VII
I	.849	-.014	.086	.040	.089	.416	-.110
II	.093	.399	.866	.042	.323	-.074	-.161
III	.338	-.178	.099	-.101	-.130	.112	-.072
IV	.093	.189	.078	.009	.115	-.052	-.281
V	.018	-.053	-.089	.047	.012	.254	.356
VI	-.074	-.078	.084	.015	-.040	.097	.551
VII	.198	.291	.042	.046	.100	.061	-.134

From the coefficient of congruence matrix, we see that Factor III of the military sample and Factor II of the private executive sample have the highest degree of agreement ( $\phi = .866$ ). These two factors were previously defined by their loadings to be concerned with taking risks involving an element of personal danger. The next highest degree of agreement between factors is between Factor I of the military sample and Factor I of the private executive sample ( $\phi = .849$ ). Both of these factors were defined by loadings on approximately the same items from the conciliation-belligerence and suspiciousness-trust scales of the PAS.

The seventh factor of the military sample and the sixth factor of the private executive sample have a moderate degree of agreement ( $\phi = .551$ ) which is apparently due to the high loading of item 35 on both factors. Also the sixth factor of the military sample and the first factor of the private executive sample have a moderate degree of agreement ( $\phi = .416$ ). As previously shown, both of these factors were defined by PAS conciliation-belligerence items.

#### D. DISCUSSION

These results suggest moderate optimism regarding the factor similarity of personality across the military and private executive samples studied. Of course this conclusion is limited in generality to the constructs assessed by the Shure-Meeker PAS.

These results, however, when coupled with those of Cummings, Harnett, and Schmidt (2), suggest caution when interpreting across samples in personality and attitudinal studies, since the number of interpretable factors, the content of factors, and the degree of differentiation within factors all differed slightly across samples. At the very least, these findings suggest

that researchers should report the factor structure of their instruments on the samples studied when comparing attitudinal responses and personality profiles across heterogeneous samples. These findings also suggest caution concerning the Mitchell and Porter (9) and Porter and Mitchell (10) comparisons of military and private executive role perceptions and need satisfactions. In those studies we have no assurance that the questionnaire used measured similar constructs in the two samples.

The emergence of three of the four *general scales* from the PAS on our two samples, as well as a similar finding in the Cummings, Harnett, and Schmidt (2) study, lends encouragement to the use of the PAS across samples. Our data do not, however, warrant an interpretation of identity of factors across samples. The differentiation within the risk scale and the prevalence of low coefficients of congruence among factors suggest that comparisons must be empirically supported in studies using the PAS (as well as other instruments) and that generalization of *complete scale content* is not warranted.

#### REFERENCES

1. CHRISTE, R., & GEIS, L. *Studies in Machiavellianism*. New York: Academic Press, 1970.
2. CUMMINGS, L. L., HARNETT, D. L., & SCHMIDT, S. M. International cross-language factor stability of personality: An analysis of the Shure and Meeker Personality/Attitude Schedule. *J. of Psychol.*, 1972, **82**, 67-84.
3. DRUCKMAN, D. Dogmatism, prenegotiation experience, and simulated group representation as determinants of dyadic behavior in a bargaining situation. *J. Personal. & Soc. Psychol.*, 1967, **6**, 279-290.
4. FRY, C. L. Personality and acquisition factors in the development of coordination strategy. *J. Personal. & Soc. Psychol.*, 1965, **2**, 403-407.
5. HARMAN, H. H. *Modern Factor Analysis* (2nd ed., rev.). Chicago: Univ. Chicago Press, 1967.
6. HARNETT, D., CUMMINGS, L., & HUGHES, G. The influence of risk-taking propensity on bargaining behavior. *Behav. Sci.*, 1968, **13**, 91-101.
7. KAHAN, J. P. Effects of level of aspiration in an experimental bargaining situation. *J. Personal. & Soc. Psychol.*, 1968, **8**, 154-159.
8. KAISER, H. F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 1958, **23**, 187-200.
9. MITCHELL, V. F., & PORTER, L. W. Comparative managerial role perceptions in military and business hierarchies. *J. Appl. Psychol.*, 1967, **51**, 449-452.
10. PORTER, L. W., & MITCHELL, V. F. Comparative study of need satisfactions in military and business hierarchies. *J. Appl. Psychol.*, 1967, **51**, 139-144.
11. SHURE, G. H., & MEEKER, R. J. A Personality/Attitude Schedule for use in experimental bargaining studies. *J. of Psychol.*, 1967, **65**, 233-252.
12. SMITH, W. P. Power structure and authoritarianism in the use of power in the triad. *J. Personal.*, 1967, **35**, 64-90.
13. VINACKE, W. E. Variables in experimental games: Toward a field theory. *Psychol. Bull.*, 1969, **71**, 293-318.



14. WRIGHTSMAN, L. Personality and attitudinal correlates of trusting and trustworthy behaviors in two-person games. *J. Personal. & Soc. Psychol.*, 1966, 4, 328-332.

*Graduate School of Business  
The University of Wisconsin  
Madison, Wisconsin 53706*

## IMPLICATIONS FOR ASSOCIATIVE PROCESSES OF SWITCHING THE MIDDLE OF THE LIST DURING SERIAL ROTE LEARNING\*

*Department of Psychology, State University of New York, College at Oswego*

---

LINDA L. BEAN AND JAMES H. McCROSKERY<sup>1</sup>

---

### SUMMARY

In an attempt to assess the degree to which specific stimulus-response associations are gradually acquired in learning a serial list, the order of the middle items was altered during acquisition. Five groups with 16 Ss per group had either no items switched, two items switched after four or eight test trials, or four items switched after four or eight test trials. The nonsense syllables were presented with slide projectors by means of standard serial anticipation procedures. Contrary to hypotheses, there were no overall differences between the four experimental groups and the control in trials to criterion or in total errors. However, although few experimental Ss reported noticing the switch, they made more errors on the trials immediately following the switch in comparison with the control group. These results are interpreted as disconfirming continuous, stimulus-specific association assumptions and supporting noncontinuous, nonassociative approaches.

### A. INTRODUCTION

There are two general approaches to the question of "what" is learned in acquiring a serial list. One approach attempts to specify the stimulus associated with each response. The classic example is the chaining hypothesis of Ebbinghaus [see Young (11)] where each item is considered the stimulus for the next item. Another example is the absolute position hypothesis which suggests that the ordinal position an item occupies on the list is the stimulus that evokes the response (2). The other approach is not concerned with specifying the stimulus for each response. In fact, Jensen (6) says that there is no specific stimulus and suggests that serial learning is the result of response integration. Slamecka (8) has suggested that the associative phase of serial

---

\* Received in the Editorial Office on December 11, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Direct requests for reprints to the second author at the address shown at the end of this article.

learning involves associating items with some nonspecific, self-generated symbol, such as first, middle, or last.

These distinctions about "what" is learned are related to questions about "how" learning takes place. Those approaches that specify the stimulus generally contend that associations are gradually strengthened between the stimulus (previous item or absolute position) and the response (*e.g.*, 5). Those approaches that do not attempt to specify the stimulus generally suggest that items are learned in some noncontinuous, maybe even all-or-none fashion (6). The purpose of the present study was to examine whether items in the middle of the list gradually acquire associative strength to specific stimuli throughout the course of list learning.

The experimental hypothesis follows from the logic of the presolution reversal technique: if items gradually acquire associative strength to specific stimuli, then altering the position of the items should retard list learning. Furthermore, it should follow that the number of items switched and the degree of practice prior to the switch should differentially affect serial list learning. In the present study the experimental groups had the order of either the two or four middle items of an eight item list altered (assumedly disrupting either three of eight or five of eight stimulus-response associations). The items were switched after either four or eight test trials. There were no changes in list order for the control group. If the S-R associations are learned in an incremental fashion, then the control group should learn the fastest, the group that had two items switched after four trials should learn not quite as fast, and the group that had four items switched after eight trials should learn at the slowest rate. The group that had four items switched after four trials and the group that had two items switched after eight trials should learn at a rate between the other two experimental groups.

## B. METHOD

### 1. *Subjects*

Eighty Introductory Psychology students were randomly divided into five groups with the restriction that there be eight males and eight females in each group.

### 2. *Materials and Apparatus*

The eight CVC syllables used with all Ss were taken from Glaze (3) with low association values ranging from .22 to .34 and averaging .28. One list order was used with the first 40 Ss, and another list order was used with the

last 40 Ss. The items that follow are in the order that all 40 of the first Ss learned to criteria: CIW, FUV, ZEC, XAK, TUJ, KIH, GAQ, and YEF. The items were individually projected on the wall by one or another of two Kodak Carousel slide projectors. Each projector had a different list order so that the alternation was accomplished without interruption for the experimental Ss by silently changing a control switch during the appropriate intertrial interval. Each carriage contained nine sets of eight slides with each set or trial separated from the others by a blank.

### 3. Procedure

The five groups were based on "amount" and "when" the middle of the list was switched. For one group, the control (Con), no items were switched. The four experimental groups comprise a  $2 \times 2$  design with two levels of amount switched (two or four middle items) and two levels of when switched (after four or eight test trials) giving the following groups: 2at4, 4at4, 2at8, and 4at8. On the basis of the control list with which all Ss ended, the starting orders were as follows for those Ss having two and four items switched, respectively: 1-2-3-5-4-6-7-8 and 1-2-4-6-3-5-7-8. The syllables were presented for 2 seconds each with an interitem interval of .7 second and an intertrial interval of 3.4 second. Standard instructions were read to each S. The Ss were required to spell the syllable before it appeared on the wall. The first list presentation was a study trial; beginning with the next trial, testing was done by the method of anticipation and continued for each S until a criterion of one errorless trial. At the end of testing, each S was asked if he noticed anything unusual about the list, and his reply was recorded.

## C. RESULTS

### 1. Overall Effects

Correct and incorrect responses were scored on each trial. According to the hypothesis, if learning occurred gradually to specific stimuli, then changing the order of the middle items should retard the learning rate according to the number of items switched and the number of trials completed before the switch. There were no significant differences between the four experimental groups ( $F_s < 1$ ) on trials to criterion. The mean trials to criterion for the four experimental groups were as follows: 23.5 (group 2at4), 24.6 (group 4at4), 24.3 (group 2at8), and 23.6 (group 4at8). As seen in the upper panel of Figure 1, the experimental groups (mean of 24.0 trials to criterion) do tend to cluster about a learning rate slower than that of the

control group (mean of 19.7). However, in the comparison of the control group with the combined experimental groups,  $F(1, 75) = 3.15$ ,  $p > .05$ , or in the comparison of the control group with each experimental group individually (all  $F_s < 1$ ), no significant differences were obtained.

The lower panel of Figure 1 compares the experimental groups with the control groups on errors at the various item positions showing the typical bowed curve. Again, since there were no differences between the four experimental groups ( $F_s < 1$ ) on total errors, they have been combined in Figure 1 for clarity. Analysis of variance of total errors (all items on all trials) per groups again showed a tendency for the control group to have fewer errors, but the difference was not significant,  $F(1, 75) = 2.27$ ,  $p > .05$ . A closer look at errors was attempted by examination of the number of errors for the control and experimental groups after the switch. The number of trial 5 to criterion errors was compared between Con, 2at4, and 4at4, but no significant difference was found,  $F(2, 45) = 1.09$ ,  $p > .20$ . The number of trial 9 to criterion errors was compared between Con, 2at8, and 4at4, but no significant difference was obtained,  $F(2, 45) = 1.07$ ,  $p > .20$ .

## 2. Immediate Effects

The above results are consistent in demonstrating the lack of an overall effect of switching the middle items of the serial list. Was there any immediate disruption caused by these alterations? The control group was compared with 2at4 and 4at4 for errors at trials 4, 5, and 6, and with 2at8 and 4at8 for errors on trials 8, 9, and 10. First look for any effects of the switch after four test trials. As can be seen in the left panel of Figure 2, there was no difference between the three comparison groups prior to the switch ( $F < 1$ ). However, a significant difference was found on the combined fifth and sixth trials,  $F(2, 45) = 4.99$ ,  $p < .05$ . A Newman-Keuls test for ordered pairs showed that the difference between Con and 2at4 was not significant, that the difference between Con and 4at4 was significant ( $p < .01$ ), and that the difference between 2at4 and 4at4 was significant ( $p < .05$ ).

Next look for the effects of the switch after eight test trials. There was no difference on trial eight (right panel of Figure 2) between these comparison groups prior to the switch ( $F < 1$ ). Immediate effects were apparent on the combined trials 9 and 10 error scores,  $F(2, 45) = 4.21$ ,  $p < .05$ . A Newman-Keuls analysis showed that the differences between Con and both 2at8 and 4at8 were significant ( $p < .05$ ) and that there was no difference between the two experimental groups ( $p > .05$ ). These two sets of findings



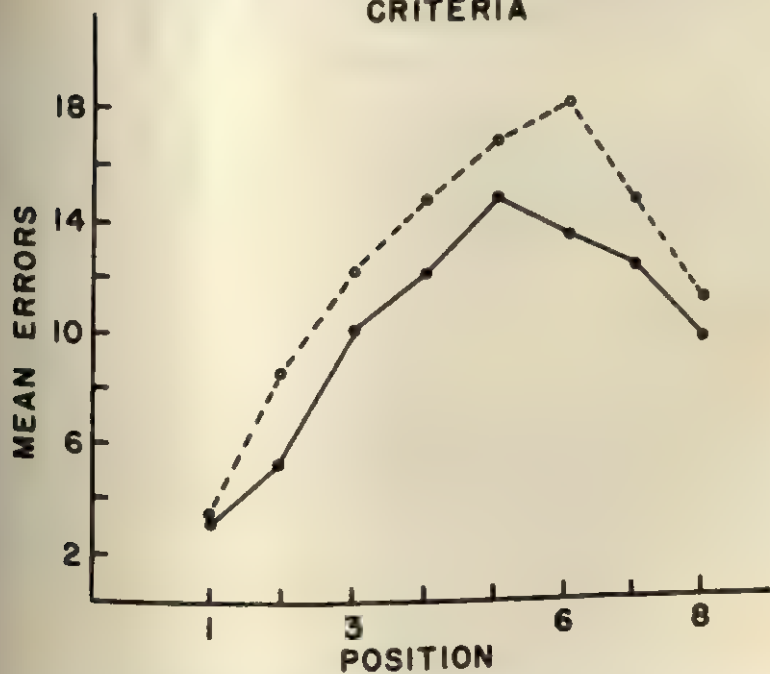
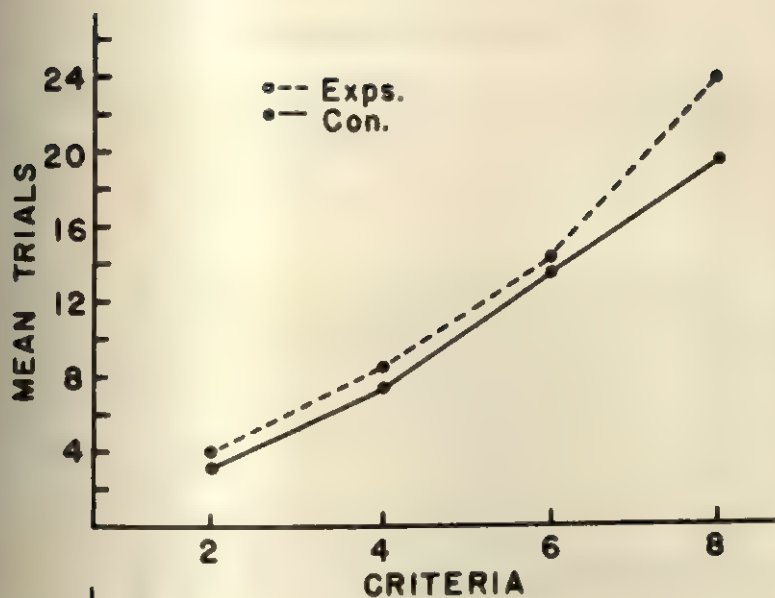


FIGURE 1

COMPARISONS ON MEAN NUMBER OF TRIALS AND ERRORS

Comparisons of the combined experimental groups and the control group on the mean number of trials on which they reached successive criteria of correct responses (upper panel) and on the mean number of total errors at the positions of the eight item list (lower panel).

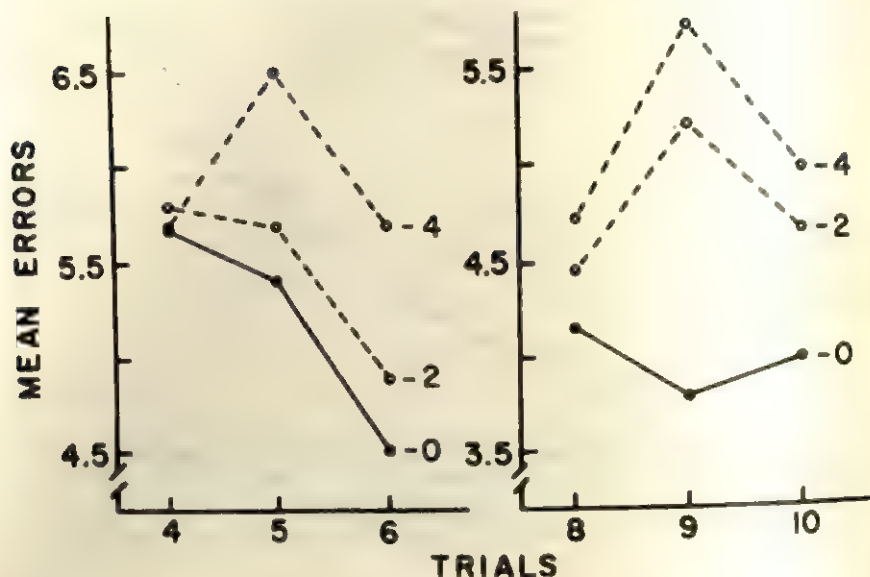


FIGURE 2

MEAN NUMBER OF ERRORS FOR SELECTED GROUPS ON SELECTED TRIALS

The left panel shows the mean number of errors on trials 4, 5, and 6 for Con (—0), 2at4 (—2), and 4at4 (—4); and the right panel shows the mean number of errors on trials 8, 9, and 10 for Con (—0), 2at8 (—2), and 4at8 (—4). Con = control; 2at4 = two middle items switched after four test trials; 4at4 = four items switched after four trials; 2at8 = two items after eight trials; 4at8 = four items after eight trials.

are consistent in demonstrating an immediate, albeit temporary, effect of altering the list.

As presented in Table 1, three categories of replies were obtained when S was asked if he noticed anything unusual about the list: (a) a "No" referred to S saying that he did not notice anything unusual about the list or making a comment irrelevant to the aspect of item order, such as "Those weren't real words"; (b) a "?" referred to S making an interrogative reply about list order, such as "was the order switched?"; and (c) a "Yes" referred to S making a declarative statement, such as "Some syllables had been switched." One control S gave a "?" reply, and the rest gave "No" replies. For the 64 experimental Ss there were 48 "No" replies, six "?" replies, and 10 "Yes" replies.

#### D. DISCUSSION

The predictions for overall learning rate based on incremental, stimulus-specific association assumptions were not confirmed. Even disrupting five of

TABLE 1

DISTRIBUTION OF RESPONSES BY Ss IN THE FIVE GROUPS TO THE FOLLOWING QUESTION:  
 "DID YOU NOTICE ANYTHING UNUSUAL ABOUT THE LIST?"<sup>a</sup>

Response	Con	2at4	Groups <sup>b</sup> 4at4	2at8	4at8
No	15	15	12	14	7
?	1	0	1	1	4
Yes	0	1	3	1	5

<sup>a</sup> See text for a description of the "No," "?," and "Yes" response categories with respect to whether Ss reported noticing the list alteration.

<sup>b</sup> Con = control; 2at4 = two middle items switched after four test trials; 4at4 = four items switched after four trials; 2at8 = two items after eight trials; 4at8 = four items after 8 trials.

eight presumed associations after nine presentations of the list did not significantly retard trials to criterion or lead to more errors as measured by serial anticipation procedures. This analysis is the same whether the specific stimulus could have been the prior item or the absolute position.

The present results provide positive replication of an earlier study by Bolles (1) where two nonadjacent middle items in a serial list of nine were switched, and a subsequent study by Jensen (7) where the two middle items in a list of 10 were switched. Bolles concluded that learning of the middle items probably does not begin until after the adjoining syllables have been learned. Bolles' noncontinuity argument was taken a step further by Jensen who suggested that each new item is added in an all-or-none fashion on a single trial.

Although the present findings are supportive of a noncontinuity theory, they do not give evidence for or against an all-or-none interpretation. This study differs from the ones previously cited by altering both the number of items switched and the number of trials prior to the switch. The values of four and eight trials were chosen on the basis of a brief pilot study. In this pilot study learning was more rapid than in the reported experiment so that, in retrospect, it might have been preferable to have switched the list after six and 12 trials. The results obtained suggest that Ss do not begin ordering the middle items until the end items have been placed. From the responses listed in Table 1 it would appear that the Ss learn very little about the order of the middle items until later in acquisition. Learning the present list might proceed about the following order 1, 2, 8, 3, 7, 4, 6, and 5. As such, if learning of the middle items occurs last, then switching the order of these items early in acquisition would have no retarding effect on overall learning. When S does begin to order the middle items, he learns rapidly, but it cannot be determined whether the learning is in large increments or in an all-or-none fashion.

Slamecka (9) has shown that Ss can learn a serial list even when the temporal input order varies from trial to trial, given that information is provided

as to the spatial location of the item. He concluded that the identification of item location is fundamental with serial learning and not the establishment of associative connections between items. More recently Horton and Turnage (4) have suggested that items from a serial list may be stored in memory in an ordered fashion, but they do not elicit one another as a chain of associations.

The temporary disruption brought about by altering the list is also consistent with these interpretations when analyzed in more detail. On the trial immediately preceding the switch only three of the 16 Ss in the 2at4 group had any of the to-be-switched items correct. On the other hand, nine of the 16 Ss in the 4at8 group had some of the to-be-switched items correct on the eighth test trial. As a result, the 2at4 group did not show a disruption, and the 4at8 group did. Some of the Ss, then, were clearly incorporating the middle items into their serial sequence when the order had to be changed. That the disruption was only temporary indicates that rather than having to unlearn three or five associations at some accumulated level of strength and then having to learn the new associations, some of the Ss only had to reorder some of the items.

In summary, the results of the present experiment along with those of Bolles (1) and Jensen (7) are interpreted as being consistent with those non-associative approaches that emphasize (a) the establishment of an ordered sequence or integrated series of independent items ("what") in (b) a non-continuous manner ("how") such that the S does not begin attempting to place the middle items into the series until the end items have been located. The present analysis is not explicit with regard to multiprocess theories (e.g., (10) which contend that serial list acquisition involves two stages (item learning and item placement). Item learning could be continuous and then the placement stage noncontinuous in the sequential manner suggested here.

#### REFERENCES

1. BOLLES, R. C. The effect of altering the middle of the list during serial learning. *Amer. J. Psychol.*, 1959, **72**, 577-580.
2. EBENHOLTZ, S. M. Serial learning: Position learning and sequential associations. *J. Exper. Psychol.*, 1963, **66**, 353-362.
3. GLAZE, J. A. The association value of non-sense syllables. *J. Genet. Psychol.*, 1928, **35**, 255-267.
4. HORTON, D. L., & TURNAGE, T. W. Serial to paired-associate learning: Utilization of serial learning. *J. Exper. Psychol.*, 1970, **84**, 88-95.
5. HULL, C. L., HOVLAND, C. I., ROSS, R. T., HALL, M., PERKINS, D. T., & FITCH, F. G. *Mathematical-Deductive Theory of Rote Learning*. New Haven: Yale Univ. Press, 1940.
6. JENSEN, A. R. An empirical theory of the serial-position effect. *J. of Psychol.*, 1962, **53**, 127-142.

7. ———. Serial rote-learning: Incremental or all-or-none? *Quart. J. Exper. Psychol.*, 1963, **15**, 27-35.
8. SLAMECKA, N. J. An inquiry into the doctrine of remote associations. *Psychol. Rev.*, 1964, **71**, 61-76.
9. ———. Serial learning and order information. *J. Exper. Psychol.*, 1967, **74**, 62-66.
10. VOSS, J. F. Serial acquisition as a function of stage of learning. *J. Exper. Psychol.*, 1969, **79**, 220-225.
11. YOUNG, R. K. Serial learning. In Dixon, T. R., & Horton, D. L. (Eds.), *Verbal Behavior and General Behavior Theory*. Englewood Cliffs, N. J.: Prentice-Hall, 1968. Pp. 122-145.

*Department of Psychology*  
*State University College*  
*Oswego, New York 13126*





## FEMINISM AND POLITICAL RADICALISM\* <sup>1</sup>

*University of Florida and University of South Florida*

---

MARGUERITE GILBERT FOWLER,<sup>2</sup> ROBERT L. FOWLER,<sup>2</sup>  
AND HANI VAN DE RIET

---

### SUMMARY

The political attitudes of 50 feminist women in relation to 'feminism' as a dimension were examined and contrasted with those of 50 of their contemporary female peers. They were administered the Attitudes Toward Feminism Belief-Pattern Scale (3), the Conservatism-Radicalism Opinionnaire (4), and a questionnaire providing biographical information and personal opinions regarding various timely political and feminine issues.

The feminist women and their peers were found to differ significantly in the attitudinal dimensions of feminism and political conservatism-radicalism. The feminist women manifested more feminism than their peers, as well as being more politically radical. Feminism as a dimension was also found to be positively correlated with political radicalism.

Both groups were also compared in their sentiments and opinions on several noteworthy issues; e.g., the potential influence of the women's vote in enhancing the status of women. Surprisingly, the feminist women and their peers failed to differ on some of the more salient of these.

In order to understand and appreciate the feminist personality, the forces potentiating the Women's movement, and the apparent similarities and differences between the feminist women and their peers, the variables of feminism, political conservatism-radicalism, and activism seem to deserve consideration.

### A. INTRODUCTION

The Women's movement has proceeded through various stages in its developmental history. From its current inception in the 1960's with small consciousness-raising groups meeting in individual homes scattered across the country, the movement has most recently culminated in well organized political

---

\* Received in the Editorial Office on December 19, 1972, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The authors' thanks are extended to Dr. Carolyn Hirsch for her valuable consultation on this study.

<sup>2</sup> Requests for reprints should be sent to the first author at the address shown at the end of this article.

<sup>3</sup> The second author is at the University of South Florida, St. Petersburg.

machinations with national platforms. There are, however, little or no well documented empirical data on the movement, and the absence of such data is particularly conspicuous in the archival journals.

The present study is concerned with the relationship between feminism and political attitudes, and is yet another attempt to examine the feminist personality and attitudes in reference to those of her contemporary peers (2). The present opportunity to study the feminist women was provided by the 1972 annual Women's Symposium of the University of Florida. This year's symposium, "Woman, Her Infinite Potential," addressed itself to political issues. It was once again well attended by several hundred women of various ages and backgrounds. It featured several prominent speakers, drawing upon academicians, politicians, and journalists. The tenor of the symposium was in keeping with the self-actualization influences of the Women's movement, and provided interested women the opportunity to discuss those political issues deemed important to the movement.

## B. METHOD

### 1. *Subjects*

The subjects (Ss) were 100 women aged 17-54 and comprised two groups. The feminist group ( $n = 50$ ) was selected from those women attending the Women's Symposium. The second group, a group of the feminists' contemporary peers ( $n = 50$ ), was drawn from classes of upper-division university women. The university women population was selected for sampling because of its anticipated similarity to the feminist group in age, educational level, and marital status. All Ss participated on a voluntary basis.

### 2. *Materials*

The materials employed were two attitude scales and a questionnaire concerning biographical information and opinions regarding feminine and political topics. The Attitudes Toward Feminist Belief-Pattern Scale (FBP), Form A (3), was used to measure "feminism" as a dimension. In particular, it is concerned with the adequacy and prerogatives of women as set forth in feminist and antifeminist propositions. The Conservatism-Radicalism (C-R) Opinionnaire, Form K (4), was used to assess political ideology. The statements cover various domains of interest, and conservatism is equated with opposition to social change, while expectation of social change is synonymous with radicalism. The questionnaire, which was constructed especially for this study, provided biographical data on the two samples of women, as well as their personal

opinions regarding various current political issues; e.g., political party affiliation, presidential candidate choice, and several timely feminine issues.

### 3. Procedure

The feminist group was obtained by asking for volunteer participants during registration, prior to the first session of the Women's Symposium in May, 1972. An envelope containing the scales and questionnaire was given to each volunteer with instructions to complete and return them by the end of the symposium, or by mail. It required approximately 30-40 minutes to complete the materials. Of 100 envelopes distributed, 66 were returned, but only 50 contained enough of the requested information to permit analysis.

The contemporary peer group was formed by asking the female members of several upper level courses in social science to participate on a voluntary basis. This group was measured within three to six weeks after the feminist group was assessed, but prior to either major political party convention.

### C. RESULTS AND DISCUSSION

The two groups were found to be not significantly different in the personal dimensions of age, educational level, and marital status ( $p > .50$ ). The mean age for the feminist women was 28.72 years,  $SD = 9.42$ , while that of their contemporary peers was 29.22 years,  $SD = 9.93$ . The modal educational level was two years of college, and single was the modal marital status.

The difference in major political party preferences between the groups was only marginally significant,  $\chi^2 = 5.75$ ,  $df = 2$ ,  $.10 > p > .05$ , with more Democrats among the feminist group and more Republicans among the peer group.

In comparing the mean feminism scores for the two groups, a highly significant difference was found ( $t = 2.91$ ,  $df = 98$ ,  $p < .01$ ), with the feminist group mean,  $\bar{X} = 39.7$ , exceeding the peer group mean,  $\bar{X} = 31.1$ . A comparison of political C-R scores between the two groups also yielded a highly reliable difference ( $t = 3.64$ ,  $df = 98$ ,  $p < .001$ ), with the peer group mean,  $\bar{X} = 19.37$ , exceeding the feminist group mean,  $\bar{X} = 14.04$ , in political conservatism.

Scores on the FBP attitude scale and on the political C-R attitude scale were correlated separately for each of the two groups. A high feminism score was found to be very reliably related to a low political conservatism (high radicalism) score; for the feminist group  $r = -.769$ ,  $p < .01$ , and for the peer group  $r = -.591$ ,  $p < .01$ . The magnitude of this negative relation-

ship between the feminism and political conservatism dimensions was not significantly different for the two groups. Feminism as a dimension is apparently positively correlated with political radicalism. The word "radicalism" is neither used nor intended as a pejorative term, but rather its meaning here is consistent with an expectation of and striving for social change and new values. This, of course, is most consistent with the purported goals of the Women's movement, which are an attempt to change the existing social structure and provide women with a more nearly equal and active vested interest in society as that of men.

The relationship between feminism and political radicalism is an intriguing one with potentially significant implications. Given the recent historical progression of the Women's movement, it seems likely that feminism is the antecedent of political radicalism or the expectation of social change. It will be recalled, however, that the feminism of the suffragettes of the early 1900's was spurred by, and actually synonymous with political radicalism, and both were embodied in primarily the women's right to vote. Since that time, as Dixon (1) notes, feminism has become a much more expansive dimension in the Women's movement, characterized by a global attitude for the movement's advocates impinging upon many domains of concern; e.g., politics, employment, marriage, education, and so on.

An analysis of presidential candidate preferences between the two groups revealed a highly significant difference,  $\chi^2 = 13.92$ ,  $df = 2$ ,  $p < .001$ . An examination of the contingency table suggests that both groups preferred McGovern to Nixon, while the feminist group also chose Chisholm over Nixon.

Both groups were asked to indicate, in a Likert-type fashion on a scale from 1-5 from "very unimportant" to "very important," how critical various issues were for them in this year's national political election. The Southeast Asian war was of uppermost importance to both groups, and the only issue similarly rank-ordered by both groups. Only the issue of inflation significantly differed in its importance for the two groups ( $t = 2.49$ ,  $df = 98$ ,  $p < .05$ ), with the peer group mean,  $\bar{X} = 4.32$ , exceeding the feminist group mean,  $\bar{X} = 3.92$ . The basic political conservatism of the peer group in comparison to the political radicalism of the feminist group may well account for this difference. It may be that the relative political radicalism of the feminist women, equivalent to expectation of social change, favors various enlightened social programs which would require fewer fiscal restrictions in order to be launched, and would, therefore, be reflected in less concern with economic inflation.



Each woman in both groups was also asked to indicate whether she thought the women's vote this year would significantly expand women's rights and opportunities, and whether or not she would cast the same vote in the 1972 presidential election as that of the primary man in her life. The lack of significant differences between the feminist women and their peers in responding to the questions concerning the women's vote this year is quite surprising and contrary to currently popular predictions (5). It seems that neither the feminist women nor their peers have any illusions about the likelihood of the women's vote in this year's election significantly enhancing the status of women, although the feminist women may be somewhat more optimistic than their peers. Equally noteworthy are the indications that only a minority of both groups contend that they would cast the same vote in this year's presidential election as that of the primary man in their lives.

All women were also asked to indicate the woman in public life whom they most admired. The feminist group nominated Shirley Chisholm as its first choice, whereas the peer group most frequently endorsed Patricia Nixon. The fact that the peer group selected Patricia Nixon as the most admired woman, while the feminist women essentially overlooked her, may suggest that the feminist women are more likely to select their role models on the basis of the model's own unique accomplishments rather than her association with an accomplished male.

To the extent that the feminist women do not differ significantly from their contemporary peers on several specific political and feminine issues, in spite of their basically different political and feminist orientations, it may be that another explanatory variable could be invoked to account for this. Perhaps as Fowler and Van De Riet (2) found previously, the feminist women and their contemporary peers are not as different as might be expected given the typical feminist stereotype. The important variable, which may account for this, may be one of "activism." The feminist women and their peers seem to share similar sentiments on many specific issues, perhaps, including several not encompassed by this study. It may be that only the feminist women are vociferous about these same issues, thereby appearing to set themselves apart from their less well identified and visible "sisters" both in their sentiments and their activism.

It would seem, then, that only when the variables of feminism, political conservatism-radicalism, and activism, are considered, can we more fully understand the feminist personality and the forces potentiating the Women's movement, as well as appreciate the observed differences between the feminist women and their contemporary peers.

## REFERENCES

1. DIXON, M. Why women's liberation-2? In Salper, R. (Ed.), *Female Liberation, History and Current Politics*. New York: Knopf, 1972.
2. FOWLER, M. G., & VAN DE RIET, H. K. Women, today and yesterday: An examination of the feminist personality. *J. of Psychol.*, 1972, **82**, 269-276.
3. KIRKPATRICK, C. The construction of a belief-pattern scale for measuring attitudes toward feminism. *J. Soc. Psychol.*, 1936, **7**, 421-437.
4. LENTZ, T. F., JR., *et al.* Manual for C-R Opinionnaire. St. Louis, Mo.: Character Res. Assoc., Dep. Educ., Washington Univ., 1935.
5. TOWARD FEMALE POWER AT THE POLLS. *Time*, March 20, 1972, pp. 33-34.

*Department of Psychology*  
*University of Florida*  
*Gainesville, Florida 32601*

## PERSONAL ADJUSTMENT OF HOSPITAL STAFF AND THEIR ATTITUDES ABOUT MENTAL ILLNESS\*

*Research Division, Central Islip State Hospital, New York*

JOHN FRACCHIA, CHARLES SHEPPARD, AND SIDNEY MERLIS

### SUMMARY

The extent to which attitudes towards mental illness and the mentally ill were influenced by the respondent's own level of personal adjustment was studied for 77 female psychiatric aides. The data suggested that, for this sample, the two variables were relatively independent of each other.

### A. INTRODUCTION

A recent study by Lieberman (3) has reported that the personal adjustment of a sample of college students was inversely related to having tolerant attitudes towards the mentally ill and to being better informed about mental illness. That is, individuals found to be the most tolerant of the mentally ill and the best informed about mental illness were themselves the least well adjusted. The author assumes these relationships are restricted to nonprofessional samples and accounts for his results by suggesting "... persons who themselves are having psychological difficulties are the ones who become interested in mental health and mental illness and inform themselves on the subject" (p. 52). Presumably, it is the increased interest and knowledge that leads to a greater expressed toleration of mental patients.

However, there are other less intellectual ways to acquire knowledge about the mentally ill. Direct experience, gained from actual contact with psychiatric patients, represents another "way of knowing" about mental illness which provides a basis for attitude formation. Iguchi and Johnson (1) found that college students who served as "volunteer" companions to psychiatric inpatients at a state hospital once a week for a semester exhibited a strong tendency to become more humanistic and less dogmatic as a result of this experience. Consequently, one might wonder whether or not individual adjustment would be an important factor affecting attitudes towards mental illness in a sample of people whose knowledge of the mentally ill derives almost entirely from daily, direct contact with mental patients: i.e., psychiatric aides.

\* Received in the Editorial Office on December 26, 1972, and published at Provincetown, Massachusetts. Copyright by The Journal Press.

One might suppose that for this type of sample, attitudes are minimally influenced by personal adjustment and maximally determined by empiric factors. If so, psychiatric aides homogeneous with respect to work experience, but differentiated as to their level of personal adjustment, should not manifest significantly different attitudes towards the mentally ill. The present study tests such a hypothesis.

## B. METHOD

Seventy-seven female psychiatric aides staffing two continued treatment services at a large state mental hospital were administered the Emotions Profile Index (EPI) and a modified form of the Opinions About Mental Illness Scale (OMI). The mean age of the subjects sampled was 45 years. They had been employed an average of 15 years and had completed 11 years of schooling.

The EPI (2) consists of paired words representing eight prototypic emotion dimensions. It is a self-report, forced choice test pairing each word with every other word on the test. The 66 items are scored on the basis of the underlying emotions they represent. The raw score is converted to a percent score which represents the number of items checked compared to the number of items which could have been checked for that emotion dimension. Scores are obtained for eight primary emotional dimensions plus the dimension of bias (a measure of social desirability): incorporation (acceptance), reproduction (joy), exploration (anticipation), protection (fear), orientation (surprise), rejection (disgust), deprivation (sadness), and destruction (anger).

The OMI (4) consists of 51 statements selected from a larger factor analyzed pool for which the responder indicates the amount of agreement on a six-point scale. The instrument is scored for five factors, as follows:

*A—Authoritarianism:* An attitude towards the mentally ill that includes authoritarian submission and anti-intracception, and views the mentally ill as an inferior class requiring coercive handling.

*B—Benevolence:* A paternalistic kindness with religious and humanistic origins rather than scientific or professionally sophisticated ones.

*C—Mental Health Ideology:* A positive philosophy reflecting current tenets of the mental health movement.

*D—Social Restrictiveness:* A view of the mental patient as posing a threat to society and the family unit; consequently, his functioning during and after hospitalization should be restricted.

*E—Interpersonal Etiology:* A cognitive rather than attitudinal factor

which reflects the belief that disturbed interpersonal experiences, such as childhood deprivation of parental love, leads to mental illness.

Two judges experienced in EPI profile interpretation separated the sample into three groups representing below average ( $N = 24$ ), average ( $N = 22$ ), and above average ( $N = 31$ ) personal adjustment according to this measure. The validity of the separation procedure by the judges was tested by analysis of variance (ANOVA) which indicated significant differences among the groups in the expected direction on the incorporation ( $F = 76.6, p < .001$ ), reproduction ( $F = 65.2, p < .001$ ), destruction ( $F = 7.1, p < .01$ ), exploration ( $F = 3.2, p < .05$ ), rejection ( $F = 8.6, p < .001$ ), and bias ( $F = 7.6, p < .01$ ) scales of the EPI. Mean differences in age, education, and duration of employment among the groups were not statistically significant. These data suggest that the three study groups were similar with respect to experience and different with respect to degree of personal adjustment.

### C. RESULTS AND DISCUSSION

ANOVA was employed to determine if any of the mean differences among the groups on the five OMI attitudinal factors were significant. Table 1 presents the mean scores of the three adjustment groups on each of the OMI factors.

TABLE 1  
MEAN OPINIONS ABOUT MENTAL ILLNESS SCALE (OMI) FACTOR SCORES  
OF THE THREE ADJUSTMENT GROUPS

OMI factor	Level of adjustment			<i>F</i>
	Low ( $N = 24$ )	Medium ( $N = 22$ )	High ( $N = 31$ )	
Authoritarianism	22	25	25	.79 (N.S.)
Benevolence	50	46	50	2.0 (N.S.)
Mental health ideology	30	31	30	.29 (N.S.)
Social restrictiveness	25	23	28	2.2 (N.S.)
Interpersonal etiology	20	19	20	.16 (N.S.)

None of the *F* ratios achieved statistical significance, indicating that for this sample, attitudes towards mental illness appeared to be relatively independent of personal adjustment as defined by EPI responses. In other words, the less well adjusted group had OMI attitude scores that resembled those of the other adjustment groups. They were no more or less tolerant of the mentally ill. They also did not espouse humanistic, professional, and etiological attitudes that were any different from the attitudes expressed by the



average and above average adjustment groups. It would appear that knowledge of mental illness acquired through actual contact and experience with a psychiatrically ill population mitigates the influence of personal self-adjustment as an attitude determinant.

#### REFERENCES

1. IGUCHI, M., & JOHNSON, R. Attitudes of students associated with participation in a mental hospital volunteer program. *J. Soc. Psychol.*, 1966, **68**, 107-111.
2. KELLERMAN, H., & PLUTCHIK, R. Manual for Emotions Profile Index. Hempstead, N.Y.: Hofstra Univ. Press, 1966.
3. LIEBERMAN, L. Attitudes toward the mentally ill, knowledge of mental illness, and personal adjustment. *Psychol. Rep.*, 1970, **26**, 47-52.
4. STRUENING, E., & COHEN, J. Factorial invariance and other psychometric characteristics of five opinions about mental illness factors. *Educ. Psychol. Meas.*, 1963, **23**, 289-298.

*Research Division*

*Central Islip State Hospital*

*Central Islip, New York 11722*

## MOTIVATIONAL CORRELATES OF PACIFIC ISLANDERS IN URBAN ENVIRONMENTS\*

*Victoria University of Wellington, New Zealand*

---

GEORGE H. HINES

---

### SUMMARY

A nonprojective measure of achievement motivation, the Lynn Questionnaire (7), was administered to 67 Western Samoans, 34 Cook Islanders, 84 Maoris, and 103 European New Zealanders in an urban work environment in New Zealand. Results indicated that Western Samoans had lower achievement motivation levels than the other three groups, Cook Islanders were lower than European New Zealanders and equivalent to Maoris, and Maoris did not vary from European New Zealanders. Findings were interpreted as supportive of previous New Zealand research and in accordance with cultural assimilation theory.

---

The use of nonprojective techniques to measure achievement motivation has received increasing attention in recent years (2, 4, 5, 7). Results from the Lynn Questionnaire have been especially promising because findings have been generally supportive of McClelland's need for achievement model (6, 8). Because of the relationship between child-rearing practices and the development of achievement motivation (1), differential levels have been found across cultures (3). New Zealand conditions offer an excellent opportunity to extend the theory, because truly multiracial elements exist side-by-side within a single economic and political structure. Among the components of this society are native Maoris, newly immigrated Western Samoans and Cook Islanders, and the numerically dominant New Zealanders of European extraction. This research was designed to determine the achievement motivation levels of these four groups as the first step in a study of socialization and adjustment.

The Lynn Questionnaire items (7) were included in a large survey of male managers and employees in Wellington and Auckland. Questions relating to occupation and ethnic origin permitted isolation of the Western Samoan, Cook Island, Maori, and European New Zealand unskilled and

---

\* Received in the Editorial Office on January 2, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

semiskilled workers whose responses are reported in this study. The questionnaires were administered at the place of employment, but were both voluntary and anonymous. The mean achievement motivation scores for the four groups were as follows: Western Samoan,  $N = 67$ ,  $\bar{X} = 2.58$ ,  $SD = 1.42$ ; Cook Islander,  $N = 34$ ,  $\bar{X} = 3.27$ ,  $SD = 1.66$ ; Maori,  $N = 84$ ,  $\bar{X} = 3.55$ ,  $SD = 1.59$ ; European New Zealander,  $N = 103$ ,  $\bar{X} = 3.96$ ,  $SD = 1.41$ .

All differences between means were tested for statistical significance by a two-tailed  $t$  test. Western Samoans were lower in achievement motivation than Cook Islanders ( $t = 2.03$ ,  $df = 99$ ,  $p < .05$ ), Maoris ( $t = 3.92$ ,  $df = 149$ ,  $p < .01$ ), and European New Zealanders ( $t = 6.21$ ,  $df = 168$ ,  $p < .01$ ). Cook Islanders scored lower than European New Zealanders ( $t = 2.12$ ,  $df = 135$ ,  $p < .05$ ), but did not vary from Maoris ( $t = .83$ ,  $df = 116$ , n. s.). Achievement motivation levels did not vary between Maoris and European New Zealanders ( $t = 1.85$ ,  $df = 185$ , n. s.).

The similarity of the Maori and European New Zealander scores supports the findings of Storm, Anthony, and Porsolt (10). Although there are differences in the child-rearing practices of the two groups (9), Storm *et al.* found that exposure to European schools and culture tended to influence positively Maori achievement motivation levels. The statistical equivalence of the motivation scores of Maoris and European New Zealanders working at the same types of jobs lends support to an interpretation based on cultural assimilation. The Western Samoans and Cook Islanders have non-European family behavior, but being newly arrived in New Zealand they have not had the opportunity to be significantly influenced by the local social environment. The scope of the present study does not permit detailed analysis of the causes of the achievement motivation differences. It does provide empirical support for previous findings and suggests potentially productive areas for future research.

#### REFERENCES

1. ATKINSON, J. W., Ed. *Motives in Fantasy, Action and Society*. Princeton, N. J.: Van Nostrand, 1958.
2. COSTELLO, C. G. Two scales to measure achievement motivation. *J. of Psychol.*, 1967, **66**, 231-235.
3. HECKHAUSEN, H. *The Anatomy of Achievement Motivation*. New York: Academic Press, 1967.
4. HERMANS, H. J. M. A questionnaire measure of achievement motivation. *J. Appl. Psychol.*, 1970, **54**, 353-363.
5. HOLMES, D. S., & TYLER, J. D. Direct versus projective measurement of achievement motivation. *J. Consult. & Clin. Psychol.*, 1968, **32**, 712-717.

6. IWAWAKI, S., & LYNN, R. Measuring achievement motivation in Japan and Great Britain. *J. Cross-Cult. Psychol.*, 1972, **3**, 219-220.
7. LYNN, R. An achievement motivation questionnaire. *Brit. J. Psychol.*, 1969, **60**, 529-534.
8. MELIKIAN, L., GINSBERG, A., CUCOLOGLU, D. M., & LYNN, R. Achievement motivation in Afghanistan, Brazil, Saudi Arabia, and Turkey. *J. Soc. Psychol.*, 1971, **83**, 183-184.
9. RITCHIE, J. E., & RITCHIE, J. Child-Rearing Practices in New Zealand. Wellington, N. Z.: Reed, 1971.
10. STORM, T., ANTHONY, W. S., & PORSOLT, R. D. Ethnic and social class differences in performance for material and nonmaterial rewards: New Zealand children. *J. Personal. & Soc. Psychol.*, 1965, **2**, 759-762.

*Victoria University of Wellington*

*P. O. Box 196*

*Wellington, New Zealand*





## ALIENATION AND IDENTITY-ROLE DIFFUSION IN LATE ADOLESCENCE\*

*Department of Psychology, The Ohio State University*

---

DOROTHY W. JACKSON

---

### SUMMARY

This study concerns the relationship between alienation and the identity and role crisis in late adolescence. Two hundred and ninety adolescents (18-21 years) divided into age sex and vocational role commitment groups were given the DJ Scale of Alienation: Self-Social Satisfaction (6). Two indexes were obtained: (a) Alienation (measured by Powerlessness and Self-Abasement) and (b) Self-Dissatisfaction (a measure of psychological disequilibrium). Powerlessness and Self-Abasement were not age/sex differentiated, whereas dissatisfaction-with-self for these attitudes was related to group differences. Young males were self-dissatisfied with their powerless status, and young females with their self-abasing attitudes. Findings were discussed in terms of the resolution of the psychological crisis and the end of the adolescent moratorium.

---

### A. INTRODUCTION

It is the purpose of this study to explore the relationship between alienation and the adolescent crisis of identity and roles. Several studies into alienation (1, 4, 7) have shown that social adaptive behavior and social and self-attitudes reflect the integration, or its lack, of the individual with the roles and mores of his society.

Although alienation is primarily a sociological study of behavior (8, 10), included in such a study are those psychosocial changes in personality and social interaction that are assumed to reflect profound changes in the social order. Yet, implicit in such research is a subjective-self focus, as hypothesized by Srole (11), which results in psychological conflict and despair. Keniston (7) and Trent and Craise (12) infer that such conflict is indicative of maladjustment and is expressed as personal dissatisfaction.

It is therefore an assumption of this study that such self-dissatisfaction

---

\* Received in the Editorial Office on January 3, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

arises from the cognitive interaction of personal values and self-percepts with an assessment by the individual of his commitment to societal role structures or their models and to the mores representing the generalized expectancies of others. According to Erikson (3) the moratorium period of adolescence ends when the adolescent has intersected his self-identities with the role structures of his society. Adolescents who are identity diffuse are unable to commit themselves satisfactorily to vocational or marriage roles, nor do they perceive fidelity in the merger of self with society.

This study was designed to determine whether two specific measures of Alienation—Powerlessness and Self-Abasement [subscales suggested by the work of Dean (2), Jackson (6), Nettler (8), Rotter (9), Seeman (10), and Srole (11)]—would indicate the disequilibrium caused by alienation from society and self at the time presumably representing the end of the adolescence moratorium.

## B. METHOD

### 1. Procedure

It is hypothesized that Self-Dissatisfaction with an awareness of Powerlessness and Self-Abasement is age and sex affected and occurs with the psychosocial conflict arising from self-role diffusion. The specific hypotheses tested are as follows: (a) There is no difference in acknowledgment of alienated-powerless feelings in adolescent males and females, but powerlessness will reflect greater self-dissatisfaction for male identities and roles than for those of females. (b) There is greater dissatisfaction due to self-abasing statements for females as compared to males, but such self-dissatisfaction decreases with age and vocational commitment.

### 2. Subjects

The sample consisted of 62 males and 228 females. Of the male group 24 ( $M_1$ ) were 18- to 19-year-olds, and 38 ( $M_2$ ) were 20- to 21-year-olds. There were 134 ( $F_1$ ) 18- to 19-year-old females, and 94 ( $F_2$ ) 20- to 21-year-old females. All subjects were in the second and third year of college. The  $M_2/F_2$  adolescents had formally made a commitment to a vocational role by declaring a teaching major.

### 3. Instrument

The instrument consisted of two subscales of Alienation—Powerlessness and Self-Abasement—derived from a six dimension DJ Scale of Alienation:

Self-Social Satisfaction<sup>1</sup> developed by Jackson (6). These two subscales represented a social assessment and a self-assessment, plus an evaluation of these responses in terms of self-standards and percepts. Powerlessness indicated the degree to which external societal factors (the roles, representatives, and generalized expectancies) had control over the consequences of one's actions and self-goals. Split-half reliability determined by Kuder-Richardson formula 8 was .88. Self-Abasement indicated the degree to which one feels self is an object of worth with self-confidence in social interaction. Split-half reliability for this scale was .75.

Item-subscale intercorrelations for DJ scales with item of the Srole-Alienation scale ranged from .60 to .72 for four of the five Srole items.

Two indexes were derived: Alienation and Self-Satisfaction. The Alienation index consisted of two scores of agreement to items indicating types of Alienation, summed separately for the two subscales under consideration. The second index—Self-Satisfaction, a measure of psychological disequilibrium—consisted of Self-Satisfaction/Self-Dissatisfaction with responses of agreement with Alienation. Of specific interest to this study are the alienated Self-Dissatisfaction responses summed through each dimension of Alienation.

### C. RESULTS

Chi squares were calculated between and within groups: (a) by sex, (b) by age for each sex group, and (c) by age/sex interaction for specific Alienation dimensions.

Neither sex nor age within sex groups contributed significantly to the acceptance of Powerlessness feelings. Approximately 1/3 of the males and females felt that situations and conditions external to them affected their actions and behavior and were beyond their control and effectiveness.

However, age did present a difference within the male group for Self-Dissatisfaction with Powerlessness. Younger males ( $M_1$ ) were more dissatisfied with themselves with this awareness than were  $M_2$  males ( $\chi^2 = 6.87$ ,  $p < .01$ ). Younger males ( $M_1$ ) were also considerably more self-dissatisfied with acknowledging external control over their behavior than either age females (for  $M_1/F_1$ ,  $\chi^2 = 26.97$ ,  $p < .001$ ; for  $M_1/F_2$ ,  $\chi^2 = 8.74$ ,  $p < .005$ ).

Self-Abasement data revealed no significant differences between sexes for self-alienation. Yet, Self-Dissatisfaction resulting from the acceptance of self-abasing attitudes was sex/age dependent. Young females were more dis-

<sup>1</sup> Scale and supportive data for subscales are available upon request from the author at the address shown at the end of this article.

satisfied with their negative self-evaluations than were older females ( $\chi^2 = 4.34$ ,  $p < .05$ ) and older males ( $\chi^2 = 8.00$ ,  $p < .005$ ), although there was no significant difference between the younger samples of either sex.

#### D. DISCUSSION

The study supports both hypotheses under consideration. Implications of the present findings provide some indication of the psychosocial crises of identity and roles during adolescence and their relationship to the multidimensional phenomena of Alienation.

Data show that older males were less self-dissatisfied with their perceptions of social powerlessness, even though there was no age difference in their acceptance of powerless attitudes. It is assumed that such awareness or acknowledgement may not, for the older males, necessarily nor sufficiently reflect lessening of personal competence and an ability to achieve in certain roles, but may be an indication of an acceptance of the limits of the adult male roles in social reality. Younger males may be more threatened by the awareness of the lack of complete self-control as masters of their own role destiny, especially as they attempt to test self-identities with the prescribed roles of manhood. The responses of the younger males may have been evidence of self-role diffusion in that they had not found a definite direction for a career goal.

There was no difference in self-criticism between the sexes, but younger females were more dissatisfied with negative self-evaluations than were the older members of either sex. This finding is congruent with the theories of Erikson (3) and Horrocks and Jackson (5) in that those individuals who can accept self-criticism have some degree of stability to their self-concepts. The greater self-dissatisfaction for younger females represents the cognitive-affective conflict experienced by their attempts at integrating negative identities with their self-concepts and values. The older groups appear to be less cognitively disequibrated and more self-adjusted despite the critical assessments of their self-worth. The older group would represent that level of psychosocial development characterized by a resolution of the identity/role crises and an end to the adolescent moratorium as indicated in part by their commitment of self to a vocational role.

#### REFERENCES

1. BAY, C. Political and apolitical students: Facts in search of theory. *J. Soc. Issues*, 1967, 23, 76-91.
2. DEAN, D. Alienation: Its meaning and measurement. *Amer. Sociolog. Rev.*, 1961, 26, 753-758.

3. ERIKSON, E. Youth, fidelity and diversity. *Daedalus*, 1962, **91**, 5-27.
4. FLACKS, R. The liberated generation: An exploration of the roots of student protest. *J. Soc. Issues*, 1967, **23**, 52-75.
5. HORROCKS, J. E., & JACKSON, D. W. *Self and Role*. Boston: Houghton Mifflin, 1972.
6. JACKSON, D. W. Self-satisfaction-dissatisfaction of older adolescents resulting from alienated attitudes. Master's thesis, The Ohio State University, Columbus, 1968.
7. KENISTON, K. The sources of student dissent. *J. Soc. Issues*, 1967, **22**, 108-137.
8. NETTLER, G. A measure of alienation. *Amer. Sociolog. Rev.*, 1957, **22**, 670-677.
9. ROTTER, J. B. Generalized expectancies for internal versus external control of reinforcements. *Psychol. Monog.*, 1966, **80**(1), Whole No. 609 (pp. 1-28).
10. SEEMAN, M. On the meaning of alienation. *Amer. Sociolog. Rev.*, 1959, **24**, 782-791.
11. SROLE, L. Social integration and certain corollaries: An exploratory study. *Amer. Sociolog. Rev.*, 1956, **21**, 709-716.
12. TRENT, J., & CRAISE, J. Commitment and conformity in the American college. *J. Soc. Issues*, 1967, **23**, 34-51.

*Department of Psychology*  
*The Ohio State University*  
*Columbus, Ohio 43210*





## TOLERANCE FOR THE UNSTABLE AND DEFENSIVE ROLE ADJUSTMENT IN RESPONSE TO SUDDEN PHYSICAL DISABILITY\*

*New York University*

---

SHELDON KASTNER

---

### SUMMARY

This study investigated the following hypothesis: Subjects tolerant of the unstable would attribute greater similarity in meaning to paired concepts referring to disabled and nondisabled roles than subjects found to be intolerant of the unstable. Thirty-eight males who had suddenly become physically disabled by traumatic injury served as subjects. The cognitive style of tolerance and intolerance of the unstable was determined by the subjects' responses to the phi phenomena. Similarity of role identification was measured by four sets of paired concepts (e.g., Me-Disabled Me) using a semantic differential format. Results of the study supported the hypothesis that subjects tolerant of the unstable are better able to integrate notions of disabled and nondisabled than intolerant subjects.

---

### A. INTRODUCTION

A sudden and dramatic shift in one's life situation often results in a state of psychological ambiguity. An individual who suddenly becomes physically disabled as a result of a traumatic injury is faced with a situation in which many of his beliefs, attitudes, and expectations are disrupted. Many of his orientations toward the world which were built upon an intact conception of self become disrupted and questioned because of the change in physical status.

Reactions to disability are strongly influenced by predisability personality patterns. Gardner, Holzman, Klein, Linton, and Spence (4) and Klein, Gardner, and Schlesinger (7) have isolated a number of persistent patterns of responding to a variety of situations. These styles of responding have been called cognitive controls. Cognitive controls refer to an individual's enduring style or strategy of information analysis or mode of reacting to

---

\* Received in the Editorial Office on January 8, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

situations. Tolerance *versus* intolerance for the unstable is one of the dimensions of cognitive controls which has been isolated (4). Individual differences in response to the phi phenomenon or apparent movement have been employed in studying this dimension. Differences in the ease of reporting apparent movement have led to studies that have found that individuals who have difficulty in reporting movement (intolerant) tend to be rigid in response to ambiguous or emotionally conflicting situations (5). The intolerant style is designed to maintain what is stable and comfortable. They (intolerants) accomplish this surface stability by focusing on outward distinctions and by avoiding depth experiences that enable emotional meaning.

The present study was designed to test the following hypothesis. Traumatically disabled subjects found to be tolerant of the unstable as compared to intolerant disabled subjects will attribute greater similarity in meaning to paired concepts referring to disabled and nondisabled roles: e.g., Me-Disabled Me.

## B. METHOD

Thirty-eight subjects were drawn from two rehabilitation hospitals in the East. All subjects were males between the ages of 16 and 30 ( $\bar{X} = 21$ ). Intelligence test data were obtained from hospital records, and only subjects within the 100-130 *IQ* ( $\bar{X} = 115$ ) range were included in the study.

A primary requirement for the study was a sudden change in the subject's way of life. The subjects selected were individuals who had suddenly become physically disabled. Prior to their injury all subjects were physically intact and had established patterns of behavior and conceptions of the world and themselves based on "nondisabled" experiences. All subjects were seen between nine months and 14 months of an accident that had led to a severe physical disability—paraplegia or quadraplegia.

The subject's response to the phi phenomenon was used as the measure of tolerance for ambiguity. As Klein (5) points out, the apparent movement or phi phenomenon creates ambiguity by evoking a discrepancy between what is known and what is perceived. Klein and Schlesinger (6) and Klein, Gardner, and Schlesinger (7) have employed the phi experience and found that subjects intolerant of phi (those reporting perceiving movement for a short period of time) have difficulty in dealing with ambiguous or emotionally toned stimuli.

The phi phenomenon apparatus consisted of (a) a motor driver electronic switching unit and (b) the stimulus box. The stimulus was divided into two sections which were painted white on the inside, each section containing a

cold cathode fluorescent lamp. Stimulus figures (two horses) cut out from black contact paper were positioned to expose two silhouettes of galloping horses. The switching unit provided a light duration of 90 milliseconds with an interval between lights that could be varied from 625 to 8 milliseconds. Subjects were run in an office where illumination was daylight restricted by venetian blinds.

Subjects were seated and given instructions. The instructions were designed to have the phi experience occur in a context of conflicting information. A brief demonstration of how the apparatus operated was followed by a statement: "As you can see nothing is really moving. The horses stay in the same position all the time. When I begin, I want you to tell me what you see."

Each subject was given five test trials, each trial proceeding from an interval of 625 milliseconds between lights to 8 millisecond intervals. Speed was increased until the subject reported that the two alternating figures now moved as a single figure (a shift from that of two horses to that of one horse moving back and forth). The alternation rate continued to increase until the subject reported two simultaneously flickering figures. The score for each trial was the range of alternation rates between the first point (movement) and second point (flicker experience) in cycles per second. The measure of tolerance for ambiguity was the subject's mean range for apparent movement for the five trials. The tolerant and intolerant groups were formed on the basis of the mean range of movement over the five trials.

### 1. *Dependent Variables*

The major dependent variable under consideration was the subject's response to major role concepts. The concepts were selected to meet two criteria: (a) relevance to the problem being studied (namely, the interest in a radical change in the subject's life—i.e., from "Male" to "Disabled Male") and (b) meaningfulness of the roles chosen for the subjects.

### 2. *Role Concepts*

The following eight role concepts were included: Me, Disabled Me, Male, Disabled Male, Father, Disabled Father, Worker, Disabled Worker. To test the subjects' attitudes toward these roles the semantic differential was employed. Nine bipolar adjectives representing the evaluative, potency, and activity dimensions were selected from a list in Osgood, Tannenbaum, and Suci (8). Both the concepts and scales were randomized to control for

positional and ordering response sets. Standard administrative instructions as suggested by Osgood *et al.* (8) were then given to each subject.

The D score, an indication of profile similarity, was computed according to the formula suggested by Osgood *et al.* (8).

### C. RESULTS

In order to test the hypothesis that subjects tolerant of the unstable would attribute greater similarity in meaning to paired concepts referring to disabled and nondisabled roles than subjects found to be intolerant, a two group multivariate analysis of variance was performed on the D scores for the four sets of concepts. D score means and standard deviations for the tolerant and intolerant groups are presented in Table 1.

A two group multivariate analysis of variance of the data revealed significant differences between the two groups over the four sets of paired concepts (Wilks lambda = .664,  $F = 4.18$ ,  $df = 4/33$ ,  $p < .01$ ). A check of the mean D scores in Table 1 indicates that, as predicted, the tolerant subjects have lower D scores across the four sets of paired concepts. Thus the hypothesis predicting greater profile similarity for the tolerant group is supported. Cooley and Lohnes (2) suggest that  $1 - \lambda$  can be used to estimate the amount of variance accounted for by the independent variable (tolerant *vs.* intolerant). In the present situation this would indicate that difference in cognitive style accounts for approximately .34 of the variance in profile similarity scores.

In order to provide additional information about the nature of the obtained difference between tolerant and intolerant subjects, a stepwise multivariate analysis of variance was performed on the profile similarity scores. The results of the stepwise analysis are contained in Table 2.

Table 2 shows that group differences in Male-Disabled Male accounted for approximately 20% of the variance. After the first variable was taken into consideration, differences in Me-Disabled Me accounted for an addi-

TABLE 1  
MEANS AND STANDARD DEVIATIONS OF D SCORES FOR TOLERANT AND  
INTOLERANT GROUPS ON FOUR SETS OF PAIRED CONCEPTS

Paired concept	Group			
	Tolerant $\bar{X}$	Tolerant $SD$	Intolerant $\bar{X}$	$SD$
Me-Disabled Me	2.80	1.52	3.90	1.39
Male-Disabled Male	3.90	1.50	5.82	2.00
Father-Disabled Father	4.66	1.84	6.53	2.58
Worker-Disabled Worker	3.76	1.80	4.65	1.74



TABLE 2  
STEPWISE MULTIVARIATE ANALYSIS OF VARIANCE

Variable entered <sup>a</sup>	F value to enter	Proportion of variance accounted	Cumulative variance
Male-Disabled Male	10.86	.198	.198
Me-Disabled Me	5.18	.132	.33
Father-Disabled Father	.04	.004	.335
Worker-Disabled Worker	.23	.001	.331

<sup>a</sup> Variables entered in order of variance accounted for in the dependent measure.

tional 13% of the variance. The remaining two variables did not significantly add to the total variance accounted for by the two previous variables.

In interpreting the data in Table 2 it should be noted that the stepwise analysis uses strictly empirical criteria for ordering the variables entered. This ordering is not based simply on their relationship with the independent variable, but also on the interrelations among the dependent measures. An interpretation of the data in Table 2 to indicate that the first two variables entered have greater theoretical importance than the remaining variable is not warranted. As can be seen in Table 1, the means for the tolerant and intolerant groups differed on all four sets of concepts. These differences ranged from one-half a standard deviation (Worker-Disabled Worker) to about one standard deviation (Male-Disabled Male). The two sets of concepts (Worker-Disabled Worker and Father-Disabled Father) which did not contribute to the overall variance reflect the fact that these variables contributed redundant information already accounted for by the variables of Me-Disabled Me and Male-Disabled Male, which were previously entered.

#### D. DISCUSSION

The results of the present study provide support for the notion that personality patterns are important factors in determining reaction to sudden onset of a disability. Subjects tolerant of the unstable were clearly better able to integrate the notions of disabled and nondisabled than intolerant subjects.

Dembo, Leviton, and Wright (3) and Barker, Wright, Myerson, and Gonick (1) suggest that the typical initial response to a physical disability, which is disruptive to the individual's personal world, is "denial of disability." The use of denial during the initial phase is not viewed as pathological. In fact, denial of the permanent implications of physical disability is seen as a necessary condition for successful rehabilitation. By avoiding the full implications of the disability the individual is better able to focus

on physical rehabilitation in an attempt to regain physical functions. Following the period of physical restitution there should begin a more realistic attitude toward disability.

A more realistic and accepting attitude toward disability requires an integration of previous "nondisabled" perceptions and beliefs with the person's changed physical status. The results of the present study indicate that intolerant subjects tend to resist this integration process. The intolerants tend to persist in denying the psychological implications of their changed physical status. They appear to compartmentalize and isolate things into separate units and classes. As Klein (5) points out, intolerants' primary concern is to maintain stability rigidly. They focus on external appearance in contrast to depth or emotional meaning. It is clear that individuals with such a rigid cognitive style would have great difficulty in integrating the psychological implications of physical disability.

An additional finding that demonstrates the pervasiveness of this cognitive attitude is the fact that differences between intolerants and tolerants were found on all four sets of paired concepts. This is again consistent with Klein's (5) notion that cognitive style operates across a wide variety of situations differing in emotional intensity.

#### REFERENCES

1. BARNER, R. G., WRIGHT, B. A., MYERSON, L., & GONICK, M. R. Adjustment to physical handicap and illness: A survey of the social psychology of physique and disability. Bulletin 55, Social Science Research Council, New York, 1953.
2. COOLEY, W., & LOHNES, P. Multivariate Data Analysis. New York: Wiley, 1971.
3. DEMBO, T., LEVITON, G. L., & WRIGHT, B. A. Adjustment to misfortune—A problem of social psychological rehabilitation. *Artific. Limbs*, 1956, 3, 4-62.
4. GARDNER, R. W., HOLZMAN, P. R., KLEIN, G. S., LINTON, H. B., & SPENCE, D. P. Cognitive Control: A Study of Individual Consistencies in Cognitive Behavior (Psychol. Issues Monog. 4). New York: Internat. Univ. Press, 1959.
5. KLEIN, G. S. Perceptions, Motives and Personality. New York: Knopf, 1970.
6. KLEIN, G. S., SCHLESINGER, H. J. Perceptual attitudes toward instability: I. Prediction of apparent movement experiences from Rorschach responses. *J. Personal.*, 1951, 19, 289-302.
7. KLEIN, G. S., GARDNER, R. W., & SCHLESINGER, H. J. Tolerance for unrealistic experiences: A study of the generality of a cognitive control. *Brit. J. Psychol.*, 1962, 53, 41-55.
8. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. The Measurement of Meaning. Urbana: Univ. Illinois Press, 1957.

Department of Educational Psychology  
New York University  
933 Commerce Building, Washington Square  
New York, New York 10003

## A PRELIMINARY INVESTIGATION OF OBSCENE LANGUAGE\*

*University of Wyoming*

---

RUSSELL FOOTE AND JACK WOODWARD

---

### SUMMARY

Both classical anthropological evidence and recent psychological research suggest the possibility that obscene language is both a linguistic universal and one of man's most frequent types of linguistic expression. The report here is of the initial results of what will be a comprehensive research project concerned with the use, function, and personal-cultural-linguistic significance of obscene language within English and in a variety of other languages. In the present study of a college student sample, an empirically derived set of linguistic obscenities was obtained, the effects of sex and production mode upon the quantity of production were assessed, and the obtained obscenities were categorized with respect to the denotative domains of experience to which they referred. In addition, a questionnaire survey of attitudes toward and use of obscenity was taken among the producing Ss. It was found that while production mode (oral *vs.* written) did not affect the quantity of obscenity produced, males significantly outproduced females. Furthermore, the obtained obscenities were meaningfully classifiable into a rather limited number of categories of social-psychological experience. Those categories containing the most exemplars seemed to reflect certain cultural attitudes toward the domains of experience represented by these categories. Finally, general analyses of questionnaire responses revealed that Ss (*a*) generally used obscenity freely, although they would limit usage around children and their own parents; (*b*) indicated restrictive reactions—particularly punishment—on the part of their parents when Ss used obscenity during their upbringing; and (*c*) gave the reason of emotional release as their primary motivation for using obscene language. Implications of the present research for future investigation—particularly across languages—were discussed.

---

\* Received in the Editorial Office on January 8, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

## A. INTRODUCTION

It is a classic anthropological fact that all societies sanction certain implicit or explicit inhibitions of behavior or linguistic expression—that is, all societies exhibit the well-known phenomenon of taboo. As has been noted by Leach (9), whether the taboo in question is behavioral or linguistic, the inhibiting sanctions are very much the same. Although the causal relationship between linguistic and behavioral taboos in the same or related domains is not always clear, linguistic taboos may well create or reinforce as well as be caused by the nonlinguistic behaviors to which they refer (14). Consequently, interest in these linguistic taboos may realistically go beyond the verbal prohibitions themselves to their possible effects upon attitudes toward the nonlinguistic behaviors to which they make reference.

A familiar manifestation of the linguistic taboo in our own society and in almost all other past and present societies (9, 10, 14, 18) is the pervasive although situationally relative prohibitions against uttering certain phonemic strings, or representing these utterances in writing, that are generally referred to as obscene, dirty, filthy, vulgar, profane, cuss, blasphemous, Anglo-Saxon, or four-letter words. It seems that with few exceptions (13) all known modern tongues contain a certain, relatively small subset of words among all the lexical items of their vocabularies that would be identified by the typical native speakers of these languages as obscene, dirty, etc. Furthermore, another aspect of the universality of this linguistic phenomenon is the fact that the obscene lexical items from those many different languages seem to refer to a rather limited number of categories of psychological and social experience. Although there seems to be differential emphasis across languages upon one or another of these categories (10), it has been suggested by Vetter (18) that verbal obscenities may be classified rather completely with respect to a quite limited number of major domains of experience: (a) words and actions relating to sex and excretion, (b) blasphemy and profanity, (c) animal abuse.

Although within our own language reference to the standard norms for frequencies of written English words (8, 17) would lead us to believe that the taboo against obscene language in English is almost completely effective, recent research by Cameron (1) suggests what we would intuitively support—that the use of obscenity is one of the more ubiquitous phenomena of *oral* English. Within the limits upon generality set by the samples used in Cameron's research, it may be tentatively concluded that "damn" is one of the 15 most frequent words in spoken English and that "hell," "God" and



"Jesus" (when either are used profanely), "fuck," and "shit" are among the 75 most often spoken words. In Cameron's study such profanities as these and others accounted for 3.5, 8.1, and 12.7% of all words uttered by speakers in on-the-job, college, and adult-at-leisure samples, respectively. It seems, then, that one particular form of linguistic taboo—that against the use of a small set of lexical items that refer to a rather limited number of experiential domains—may be tentatively referred to as a linguistic universal [see Greenberg (5)], one that has not been previously noted as such in the search for regularities across languages.

Given the essential universality of obscenity across languages—both in terms of the existence of the phenomenon and the common referents involved—as well as the frequency with which it is used (at least in American English), it is somewhat surprising that psycholinguistic research has been essentially unconcerned with the phenomenon, particularly during the past several years when the implicit sanctions against research into previously taboo topics appear to have been relaxed. However, apart from Cameron's work (1) and a few investigations of written obscenity or graffiti (e.g., 16), research concerned with obscene language has been nonexistent except incidentally when researchers have employed obscene words as stimuli when investigating such phenomena as perceptual defense (2) and personality styles (4, 15).

In view of what has been said above regarding the ubiquitous taboo against obscenity, the equally widespread violation of this taboo in our own language, and the possible effects of these taboos upon nonlinguistic behavior, it would appear that research concerned with obscene language may prove fruitful with respect to providing insights into (a) linguistic (particularly semantic) universals and (b) the psychology and sociology of the user of obscene language and his less inclined counterpart. Furthermore, it would seem that any behavior as widespread as the use of obscene language is an intrinsically appropriate research domain for those interested in a complete understanding of human behavior and its motivation.

Most generally, then, the present research was undertaken as the beginning of a long-range project aimed at rectifying the minimal attention previously given to the linguistic phenomenon of obscenity. The immediate goals of the present investigation, however, were fourfold: (a) to obtain an empirically derived set of obscene words (as opposed to an experimenter-imposed set) for use in future research; (b) to identify the categories of experience that such empirically derived obscenities refer to; (c) to investigate possible differences in the production of obscenity as a function of the mode (oral *vs.*



written) of production and the sex of the producing Ss; and (*d*) to make a preliminary assessment of the attitudes toward and previous life experience with the use of obscene language. The first goal, in addition to providing a standard set of stimulus items for subsequent research, was also related to establishing a tentative ordering of the relative frequencies of obscenities in English when individuals are asked to produce these particular lexical items freely. The second goal was concerned with relating obscenity to nonlinguistic categories of human experience. The third goal followed from previous research that has suggested that production mode and sex of the producing S (18) are variables that may influence verbal production. With respect to assessing the role of sex differences in the production of obscene language, special care was taken (see Method section) to control for social inhibitions upon production that may have differential consequences across the sexes. Finally, the preliminary attempts to assess Ss' attitudes toward the use of obscenity and their previous life experience with obscene language—particularly during their childhood years—were undertaken to provide a basis for the evaluation of personality and social influences in the use of and reactions to obscene language.

## B. METHOD

### 1. Subjects

The Ss were 40 (20 males and 20 females) students enrolled in introductory psychology courses at the University of Wyoming. All Ss were volunteers, and their median age was 19.2 years. It should be noted at this point that the authors are aware of the restrictions upon generalizing from this sample because of the rather narrow age-range and select educational background of its members.

### 2. Materials and Procedure

Ss were randomly assigned to two groups, A and B, so that there were 10 males and 10 females in each group. When each S arrived at the experimental room, he was advised by E that he would be involved in an experiment that would be relatively self-controlled. That is, each S was ushered into the experimental room, seated at a table, and provided with a type-written page of instructions that he was advised would be self-explanatory. E then left S alone in the experimental room. Ss in Group A found the following instructions along with a pencil and paper on the table at which they were seated:

"You are taking part in an experiment which is designed to determine what words are considered obscene by a significant part of the population. To say a word is obscene for you doesn't mean it is obscene to me, and *vice versa*. In this experiment, we want you to write a list of all the words or combination of words that *you* think are obscene (that is, dirty, vulgar, foul, or generally objectionable) to a significant part of the population. As stated before, not all words are obscene to all people. A word that might be very offensive to the elementary school teacher could possibly be part of a dock-worker's everyday language. But on this list, we want you to write *all* the words that are obscene to you, and that you think would be obscene to other people, generally speaking. If you think of a word and consider it to be mildly obscene, please list it also. And if you think of a word that you consider terribly obscene, please list it too. The point is to produce as many words as you can that you think at least some significant number of people in the population would consider obscene. Because of the nature of this experiment, and what might seem to you a strange choice of subject matter, we would like to impress upon you the fact that this is a genuine, legitimate psychological experiment, and we would like you to make a serious attempt at listing the words per instructions so that our data will be as accurate as possible. Before we procede any further, if you feel that this experiment will be offensive to you in any way whatsoever, please come out of the lab and inform the experimenter that you don't want to take part in this experiment. Thank you."

Ss in Group B found, instead of paper and pencil, a Craig cassette tape recorder, Model 6200 on the table and a page of instructions that differed from those presented to Group A only in terms of referring to producing the words into the tape recorder and giving simple directions on how to activate the tape recorder by pressing an "on" switch on the recorder's hand microphone. The emphasis upon allowing Ss complete physical privacy while they were producing their obscenities was, of course, to minimize any potential effects of the presence of other people upon production. [See Vetter (18) for a review of experimenter effects upon expression of taboo words.] It should be mentioned, incidentally, that all but one S who volunteered for the experiment—the purpose of which Ss were unaware until they read the experimental instructions—completed the task asked of them. Although all Ss were given unlimited time to produce their obscenities, most Ss in both production mode groups finished within less than two minutes.

As each S left the experimental room upon completing the production task, he was taken to a private adjoining room, presented with a 21-item question-

naire, and asked to fill it out anonymously at his leisure and return it to the experimenter. Questions generally dealt with such topics as Ss' (a) frequency of usage of obscenity in speaking and thinking; (b) attitudes toward the use of obscenity by and around others, in general, and in various situations; (c) experience with obscenity in their upbringing and home life; (d) parents' reactions to Ss' use of obscenity as children and presently with their parents; (e) opinions as to what words or phrases seemed most obscene; (f) differentiation or lack of it between such words as "obscenity," "profanity," "vulgarity," "blasphemy," etc.; and (g) reasons why they used obscenity in speech. The format of questions varied among those amenable to "yes," "no," or "sometimes" answers; ratings on a five-point scale; and open-ended questions.

### C. RESULTS

All recognizable words, word combinations, or phrases—with the exception of repeats—that Ss produced in writing or into the tape recorder were counted as single obscene responses. Particularly idiosyncratic responses were included, although such responses were relatively rare. Recognizable obscenities from foreign languages, which were similarly rare, were also counted.

#### 1. Production of Obscenities

Table 1 shows the means and standard deviations for the number of obscene responses produced as a function of production mode (oral *vs.* written) and sex of the producing S. With use of these variables as factors, it was found through a two-factor analysis of variance that while production mode made no significant difference in the number of obscenities produced ( $F = .93$ ,  $df = 1/1$ ,  $p > .05$ ), males significantly outproduced females ( $F = 4.58$ ,  $df = 1/1$ ,  $p < .05$ ) by a magnitude of approximately 50%. No significant interaction effect ( $F = 1.13$ ,  $df = 1/36$ ,  $p > .05$ ) was revealed by the analysis.

TABLE 1  
MEANS AND STANDARD DEVIATIONS FOR NUMBER OF OBSCENE WORDS  
PRODUCED BY SEX AND BY PRODUCTION METHOD

Production mode	Male		Female		Total	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Oral	20.7	9.0	11.5	7.1	16.1	9.3
Written	16.4	11.2	13.3	6.3	14.9	9.2
Total	18.6	10.4	12.4	6.8	15.5	9.3

Table 2 presents the 20 most frequently produced obscenities by the 40 Ss. These were obscenities produced by at least 20% of all Ss. Since no differential effect as a function of production mode or sex of S was noted with respect to these frequencies, Table 2 does not include these distinctions. However, Table 2 does indicate the primary denotative referent type(s) for each of the 20 most frequently produced obscenities. While the exact labels used to indicate these referents (i.e., Ancestral Allusion, Body Process, etc.) are arbitrary phraseology of the authors, the denotative classification of each of the 20 obscenities was guided by reference to the *Dictionary of American Slang* (19) and seems to have face validity.

Table 3 presents the total number of obscene responses produced by denotative referential class. Table 3 was prepared in the following manner. The authors, working as independent raters, divided all obscenities produced by all Ss into any one or more of the eight classes of denotative reference indicated in Table 3. These eight classes were decided upon as a consequence of previous discussion between the authors as to the most obviously occurring domains of denotative experience referred to by all of the obscenities. As with the classifications indicated in Table 2, the *Dictionary of American Slang* (19) was frequently consulted. When both authors independently concurred regarding the classification of an obscenity into

TABLE 2  
MOST FREQUENTLY PRODUCED OBSCENITIES BY 40 SUBJECTS  
REGARDLESS OF SEX OR PRODUCTION METHOD

Word	Frequency	Denotative classification
Fuck	33	Body Process
Shit	33	Body Product
Bastard	25	Ancestral Allusion
Cunt	23	Body Part
Motherfucker	20	Body Process/Ancestral Allusion
Cocksucker	19	Body Process
Son-of-a-Bitch	19	Animal/Ancestral Allusion
Bitch	18	Animal
God Damn	15	Religious Blasphemy
Prick	15	Body Part
Damn	12	Religious Blasphemy
Whore	10	Social Deviation
Hell	10	Religious Blasphemy
Asshole	10	Body Part
Cock	9	Body Part/Animal
Piss	9	Body Product
Tit(s)	9	Body Part
Suck	8	Body Process
Bullshit	8	Animal/Body Product
Nigger	8	Ethnic-Racial Slur



one or more classes, it was categorized accordingly. In those instances when the authors disagreed or were unable to classify an S's obscenity into any of the first eight categories, the obscenity was rated as unclassified. While most of the classes indicated in Table 3 are self-explanatory, it should be pointed out that those obscenities that have both a substantive and a verb form (i.e., "shit") that refers to essentially the same domain of experience (i.e., excretion) were classified only with respect to their substantive form. Thus "shit" is classified only as a Body Product and not as a Body Process as well. However, obscenities that have multiple denotative meaning in their substantive or in any form alone were classified into more than one category. For example, "son-of-a-bitch" was categorized as both Animal and Ancestral Allusion. With respect to all such classification, however, it must be emphasized strongly that while the rationale behind the classification was to indicate the prevailing denotative domains of experience referred to by the various obscenities produced, it is recognized that in actual usage the meaning of obscenity is much more related to expletive, emotional release through epithet and is probably rarely connected to the denotative referential domain of the obscenity being used. It is unlikely that when one is called a "bastard," for example, that one would seriously consider that the individual addressing him had evidence regarding the legitimacy of his birth.

## 2. *Some General Impressions from Questionnaire Responses*

Since the primary purpose for seeking information requested in the questionnaire items was to provide a more intelligent basis for future research

TABLE 3  
TOTAL NUMBER OF OBSCENE WORDS PRODUCED BY DENOTATIVE CLASS  
REGARDLESS OF SEX OR PRODUCTION METHOD

Denotative class	Exemplars produced	Most frequent exemplar
Body Process	196	Fuck (33)
Body Part	127	Cunt (23)
Body Product	60	Shit (33)
Ancestral Allusion	85	Bastard (25)
Religious Blasphemy	60	God Damn (15)
Animal	50	Son-of-a-Bitch (19)
Social Deviation	40	Whore (10)
Ethnic-Racial Slur	18	Nigger (8)
Unclassified	35	Fascist (2) Fuckin' (2)

*Note:* Some obscenities were assigned to two or more denotative classes. Numbers within parentheses are frequencies.



into developmental, sociological, and psychological factors in the use of and attitudes toward obscenity, general rather than detailed analyses of the questionnaire responses were undertaken. A number of general findings based upon all 40 Ss' answers will be reported here. To begin with, almost all Ss admitted to using obscenity in speech or thought with some degree of regularity. All but eight Ss indicated that they used obscenity in speech at least sometimes, and 25 of the 40 Ss indicated a similar degree of usage of obscenity in thought. In fact only one S indicated that she never used obscenity in speaking or thinking. There was no discernible sex difference with respect to such reported usage.

Secondly, while few of the Ss indicated that they objected to the use of obscene language in general, the majority indicated that they would restrict its use around young children. To the general question "Do you object to the use of obscenity?" 20 Ss indicated that they would rarely or never object, 15 stated that they sometimes would, and no S indicated that he always would. Again, there was no apparent sex difference in responding. However, when Ss were asked about their objections toward the use of obscenity in specific contexts (such as a public place, in the home, etc.) nearly one-half, as stated above, of all Ss (with no apparent sex difference) suggested that one should refrain from using obscenity in the presence of children. Furthermore, over one-half (24 of the Ss) reported that they objected to children themselves using obscene language. On the other hand, over three-fourths (32 of the Ss) indicated that they rarely or never objected to their friends speaking obscenely. Similarly, 33 of 40 Ss responded that they had no objections to the use of obscenity by members of the opposite sex. Interestingly, six of the seven Ss who said that they objected were males. With respect to a final situational question, exactly one-half of the 40 Ss said that they did not condone the use of obscene speech in a public place. All but five of these objections were made by females.

Thirdly, Ss' responses to those questionnaire items dealing with the use of and reactions to obscenity in their previous home life and in present relationships with their parents indicated that 25 of all Ss reported little or no use of obscenity in their home while growing up, and that only about one S in five reported extensive use of obscenity in his early home background. Furthermore, questionnaire analyses indicated that by far the most common reaction by parents to the use of obscenity by the Ss as children was some sort of verbal or physical punishment, or both. Seventeen Ss indicated that they were generally verbally reprimanded by their parents for using obscenity, seven Ss cited physical punishment as most prevalent, and

two Ss reported both kinds of punishment as common. Several Ss reported that one or both parents reacted to obscene language by their offspring with emotional tirades. In fact, only six Ss reported parental responses that could be conceived of as predominantly rational and objective. It was not unexpected, then, to find that 25 of 40 Ss reported that they rarely or never use obscenity in their parents' presence now, with only one S reporting that he presently uses obscenity frequently with his parents.

Fourth, questionnaire responses suggested that overwhelmingly obscenities making denotative reference to either Body Parts, Products, or Processes were the obscenities from which Ss selected "the most obscene word" of which they could think. Words from these three denotative categories accounted for 35 of the 40 Ss' choices as "the most obscene" word, with "fuck" being the choice 24 times.

Fifth, with respect to Ss' opinions as to what distinguishes words they considered profane or blasphemous from those regarded as dirty, vulgar, or obscene, it was generally apparent that Ss could not readily make consistent distinctions or made no effort to do so. Finally, all but one of 40 Ss gave some variation of the reason of emotional release (i.e., "express my feelings," "let off steam," "because I feel mad") as their most common motive for using obscenity.

## D. DISCUSSION

### 1. *Frequencies and Denotative Referents of Obscenities*

In the present research the obscenities produced by the 40 Ss—regardless of sex or production mode—were given with differential frequencies that are comparable to those obtained by Fiedler (3) in an unpublished study requiring Ss to produce spontaneously as many obscenities as they could. Although all Ss in Fiedler's study were males, "fuck" and "shit," as in the present study, were the most often given obscenities. Similarly, in Cameron's research (1) into observed, everyday speech—although "damn" was the single most frequently uttered obscenity—"fuck" and "shit" were of relatively high frequency.

Furthermore, the relative frequencies of obscenities from various denotative domains were comparable between the present study and Cameron's (1) investigation. Although Cameron classified the obscenities produced in everyday speech by college students at leisure into different categories from those arrived at in the present study, when the obscenities that were tabulated in Cameron's investigation are reclassified in accord with the categories used

in the present research, it is seen that in both studies obscenities referring to Body Processes are more frequent than those referring to Body Parts, which in turn are more frequent than those referring to Body Products. Also in both studies, obscenities concerned with Social Deviation or Ethnic-Racial Slurs are of lower frequency than obscenities with other denotative referents. The only striking difference between Cameron's findings and those of the present study is that in Cameron's research obscenities denoting Religious Blasphemy were the most frequent obscene words produced. Since the obscenities tabulated by Cameron were obtained in the context of public discourse with its usual social restraints, it is not surprising to find that the most frequently uttered obscenities were those which are probably now the least taboo in modern spoken English (14). Furthermore, it is not surprising that when the restraints and taboos of social discourse are essentially eliminated by procedures incorporated into the part of the present study that was concerned with spontaneous, obscene production, those obscenities regarded as perhaps most offensive would be most often expressed, and those having milder taboo value (i.e., Religious Blasphemy) would find less expression.

That those obscenities which refer to bodily products, parts, and processes that relate to sexuality, excretion, or both were the most commonly produced in the present investigation seems a rather direct reflection of certain of the cultural values of the producing Ss. Even allowing for the fact that some obscenities produced reflected more than one denotative class, approximately 60% of all obscenities produced had as their only denotative referent either a human body part, product, or process. Furthermore, the *first* obscenity produced by 38 of the 40 Ss represented one of these three denotative realms. It was not surprising, then, when Ss were asked on the questionnaire what word seemed "most obscene" to them that all but five Ss indicated a word reflecting one of these three bodily related categories. It has been widely supposed in literary and anthropological accounts of so-called dirty language that although the number of denotative realms of experience referred to by the majority of languages is quite limited, there is differential emphasis across languages as to what categories of experience are emphasized in terms of providing speakers of the languages with obscene lexical items. Two of the most thorough accounts along these lines—by Sagarin (14) and Legman (10)—have speculated that for speakers of American English, obscenity and its various manifestations (particularly in expletive, anecdote, and humor) are primarily grounded in references to taboo features of the human body, usually in the sexual or scatological sense. The data noted above seem to corroborate their speculations.

## 2. *The Effects of Sex and Production Mode*

As expected, the sex of the producing Ss significantly affected the quantity, though not the quality, of production, with males outdoing their female counterparts by 50%. Previous research has suggested the potential superiority of males in any task akin to a cued but otherwise unrestrained output of obscenity (6, 7, 11, 12). However, more confidence may be placed in the difference noted in the present study to the extent that previously unused controls were invoked against possible differential effects across sexes of social inhibitions in the immediate physical environment. Being in a better position to accept that the sex difference in quantity of production is a real one, we are left with the task of trying to account for it. The most ready answer, of course, would seem to lie in differences in acquired male and female sex roles and their expected behaviors associated with these roles. However, since our questionnaire data and the quality (kinds) of obscenities produced did not suggest sex differences, it probably best remains for future investigation to unravel the relationship—if any—between differential upbringing as a function of sex and the tendency to produce obscenity—in the laboratory or the real world.

With respect to the lack of effect of production mode upon at least the quantity of obscenities produced, it seems reasonable to expect that when the interaction between situational social restraints and mode of expression (writing *vs.* speaking) is reduced through minimizing situational social inhibitions, the effects of mode of expression upon production of obscenity would be correspondingly minimized. In a socially inhibiting situation, one may feel more comfortable expressing taboo words in the more private vernacular of writing. However, in a situation where social inhibitions are at a minimum, one is just as apt to feel safe in producing socially taboo forms orally as he is graphemically.

Although there was no significant interaction between sex and production mode, it is noteworthy that females tended to produce more obscenities in written form and were closer to matching males in this mode, while males tended to be more productive orally and nearly doubled the output of females when called upon to speak obscenities. These tendencies suggest that while performance of the Ss overall did not differ as a function of production mode (because of the general absence of social inhibitions in the experimental setting), females nevertheless brought sufficient self-generated inhibitions (perhaps as a function of sex role) so that they retained some tendencies to



inhibit their expression of obscenity through the more blatant oral expressive mode.

## REFERENCES

1. CAMERON, P. Frequency and kinds of words in various social settings, or what the hell's going on here. *Pacific Sociolog. Rev.*, 1969, Fall, 101-104.
2. DIXON, N. Subliminal Perception. New York: McGraw-Hill, 1971.
3. FIEDLER, T. The componential meaning of dirty words. Unpublished Master's thesis, University of Tennessee, Knoxville, 1970.
4. GALBRAITH, G., HAHN, H., & LEIBERMAN, H. Personality correlates of free-association sex responses to double-entendre words. *J. Consult. & Clin. Psychol.*, 1968, **32**, 193-197.
5. GREENBERG, J., Ed. Universals of Language. Cambridge, Mass.: M. I. T. Press, 1966.
6. GROSSER, G., & WALSH, A. A. Sex differences in differential recall of taboo and neutral words. *J. of Psychol.*, 1966, **63**, 219-227.
7. HUNTER, E., & GAINES, B. Verbal taboo in a college community. *Amer. Speech*, 1938, **13**, 97-107.
8. KUCERA, H., & FRANCIS, W. Computational Analysis of Present-Day American English. Providence, R. I.: Brown Univ., 1967.
9. LEACH, E. Anthropological aspects of language: Animal categories and verbal abuse. In E. Lenneberg (Ed.), *New Directions in the Study of Language*. Cambridge, Mass.: M. I. T. Press, 1966.
10. LEGMAN, G. The Rationale of the Dirty Joke. New York: Grove Press, 1968.
11. MCGINNES, E. Emotionality and perceptual defense. *Psychol. Rev.* 1949, **56**, 244-251.
12. MILLER, M., & SOLKOFF, N. Effects of mode of response and sex of E upon recognition thresholds of taboo words. *Percept. & Motor Skills*, 1965, **20**, 513-578.
13. PEI, M. The Story of Language. New York: Mentor, 1960.
14. SAGARIN, E. The Anatomy of Dirty Words. New York: Stuart, 1962.
15. SCHILL, T., EMANUEL, G., PEDERSON, V., SCHNEIDER, L., & WACHOWIAK, D. Sexual responsivity of defensive and nondefensive sensitizers and repressors. *J. Consult. & Clin. Psychol.*, 1970, **35**, 38-44.
16. SECHREST, L., & FLORES, L. Homosexuality in the Philippines and the United States: The handwriting on the wall. *J. Soc. Psychol.*, 1969, **79**, 3-12.
17. THORNDIKE, E., & LORGE, I. The Teacher's Word Book of 30,000 Words New York: Columbia Univ., Teach. Coll. Press, 1944.
18. VETTER, H. Language Behavior and Communication. Itasca, Ill.: Peacock, 1969.
19. WENTWORTH, H., & FLEXNER, S. Dictionary of American Slang. New York: Crowell, 1967.

Department of Psychology  
University of Wyoming  
P. O. Box 3415, University Station  
Laramie, Wyoming 82070





## SPEED OF MOTOR CONFLICT RESOLUTION AS RELATED TO TYPE OF CONFLICT AND MANIFEST ANXIETY\* 1

*University of Richmond*

---

JEFFREY C. FRACHER<sup>2</sup> AND KENNETH A. BLICK<sup>3</sup>

---

### SUMMARY

In a study of motor conflict resolution, 20 Ss in each of two groups, designated as high drive (HD) and low drive (LD) according to scores on the Taylor Manifest Anxiety Scale, had to resolve approach-approach (AP-AP), avoidance-avoidance (AV-AV), and double approach-avoidance (DAP-AV) conflicts. The HD group took significantly longer to resolve all three conflicts than did the LD group; however, resolution time for both drive groups did not differ for the AV-AV and DAP-AV conflicts, but was significantly faster for the AP-AP conflict. It was concluded that manifest anxiety was inversely related to performance in complex conflict situations, and several areas of follow-up research were indicated.

---

### A. INTRODUCTION

Lewin (5) is primarily credited with systematically introducing the topic of conflict into psychology. According to Lewin's field theory (6), an organism in a conflict situation experiences tension which results in restless nondirected behavior. Objects in the organism's environment give the restless behavior direction as a result of what is known as valence. If an object is attractive to the organism, it is said to have positive valence, whereas if an object is repelling, it is said to have negative valence. Therefore, positive valences elicit approach responses in the organism, and negative valences elicit avoidance responses.

Using his concepts of valences, vectors, and field forces, Lewin (6) identified three types of motivational conflicts. In a Type I conflict the organism is

---

\* Received in the Editorial Office on January 11, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was supported in part by a Faculty Research Grant from the University of Richmond.

<sup>2</sup> Now at Western State Hospital, Staunton, Virginia.

<sup>3</sup> Requests for reprints should be sent to the junior author at the address shown at the end of this article.

attracted to two positive valence objects. If the organism is simultaneously confronted with a positive valence object and a negative valence object in the same field, a Type II conflict is said to exist. A Type III conflict involves having to choose to avoid one of two objects with negative valence.

Hovland and Sears (2) extended Lewin's conceptualizations to include a Type IV conflict which consists of two Type II conflicts occurring together. In other words, the organism has to choose between one of two goals, each of which has both positive and negative valences. They identified the conflicts as: Type I—approach-approach (AP-AP); Type II—approach-avoidance (AP-AV); Type III—avoidance-avoidance (AV-AV); and Type IV—double approach-avoidance (DAP-AV). A Type IV conflict was thought by Hovland and Sears (2) to best approximate real life conflicts.

One of the best known experimental attempts to produce conflict behavior was a study of motor conflict by Hovland and Sears (2). Their apparatus consisted of a pencil and a metal board covered with paper and equipped with four lights, two red and two green. The lights formed a system for simulating the four types of conflict. For example, in the AP-AP conflict *S* was told to draw a line to the green light that flashed. After 20 trials in which one light was flashed at a time, a green light at both corners flashed, and *S* was faced with a conflict. The other types of conflict were simulated by various combinations of instructions and flashing lights. The authors were primarily concerned with the degree of difficulty encountered with the resolution of each conflict type and the mode of resolution most frequently utilized for each conflict type. The following four modes of conflict resolution were available to the *Ss* in the motor conflict situations: (*a*) a single response meant choosing one of the lights and responding solely to it; (*b*) a double response meant that *S* attempted to satisfy both demands and drew a line first to one light and then to the other; (*c*) a compromise response satisfied neither demand completely, since a line was drawn to the center of the lights; and (*d*) blocking was designated when *S* failed to move the pencil from the starting position.

The results of the study indicated that the AP-AP conflict was the most easily resolved, since a clear majority gave nonconflict single responses. The AP-AV and AV-AV conflicts were typically resolved with double and blocking responses, respectively, and were therefore judged more difficult to resolve than the AP-AP conflict. The DAP-AV conflict was found to be the most difficult to resolve because blocking occurred in almost three-fourths of the cases. Lewin's observations that each conflict was associated with specific modes of reaction were supported.

The Taylor Manifest Anxiety Scale (MAS) developed by Taylor (9) has received widespread usage in psychological research as a psychometric measure of generalized drive with the majority of experimental applications concerned with the role of drive in performance of a task. Drive level can be varied by means of selection of Ss on the basis of extreme scores on the MAS, rather than by such experimental manipulation as electric shock or stress-producing instructions.

The first study to utilize the MAS was reported by Taylor (8) and involved the conditioned eyelid response. Using one group each of high drive (HD) and low drive (LD) Ss, Taylor presented an airpuff to the S's right eye as the unconditioned stimulus, following a conditioned stimulus which was an increase in brightness of an illuminated disc. As measured by the percentage of conditioned responses (eyeblink responses) and trials to extinction of the conditioned response, the HD group was clearly superior in the amount of conditioning to the LD group. Taylor (10) interpreted these results to indicate that MAS scores reflect differences in a chronic emotional state so that Ss scoring high on the scale tend to bring with them a higher level of emotionality to the experimental situation than do Ss scoring at lower levels. In other words, differences between HD and LD groups should be found with use of the MAS whether or not there is a threat present in the form of noxious stimulation, fear of failure, etc.

After training Ss on a key pressing response, Wenar (12) measured the reaction time of HD and LD Ss to three different stimuli presented in varying degrees of intensity—a buzzer, a weak shock, and a strong shock. The results indicated that reaction time was significantly related to both drive level and stimulus intensity, with response time being quicker as these variables increased.

It would appear from the studies discussed above that in simple conditioning experiments, HD Ss tend to demonstrate superior performance than do LD Ss. Other studies have shown however, that as the experimental task increased in complexity, the performance of LD Ss surpassed that of HD Ss (3, 11). For example, Taylor and Spence (11) found that HD Ss required a greater number of trials to reach criterion in a verbal learning situation involving competing responses than did LD Ss. On the basis of numerous lines of evidence, Child (1) concluded, with regard to the MAS, that as "the task becomes more complex (in the sense of involving conflict among various response tendencies) there is a tendency for high anxiety Ss to show increasingly poor performance in comparison with low anxiety Ss" (p. 154).

In view of the above findings concerning conflict resolution and manifest

anxiety, the present study attempted to blend the separate areas of research by investigating the effect of drive level on resolution time of AP-AP, AV-AV, and DAP-AV motor conflicts. If performance on complex tasks was inversely related to MAS scores, HD Ss would have longer resolution times for all three conflicts; and resolution times for both LD and HD Ss would increase as the conflicts increased in complexity.

## B. METHOD

### 1. Subjects

A total of 124 students from introductory psychology classes at the University of Richmond were given the Taylor MAS as a preliminary screening device. Selection of groups was based upon the procedure recommended by Taylor (9) for use in studies employing the MAS to define operationally drive levels in human Ss. Two groups of 20 Ss were chosen on the basis of extreme scores on the MAS, and those students whose scores were in the upper 15% of those tested on the MAS were placed in the HD group. The HD group contained 14 males and six females whose scores ranged from 28 to 43 anxiety responses out of a possible 50, with a mean score of 35.8. Those students whose MAS scores were in the lowest 15% of those tested were placed in the LD group, which also contained 14 males and six females; their scores ranged from 1 to 11 anxiety responses with a mean of 6.2.

### 2. Apparatus

The apparatus was a variation of the motor conflict board designed by Hovland and Sears (2). It consisted of a plywood base, three feet long and two feet wide, which was divided in the middle by a partition 18 inches high. Located on *S*'s side of the board were four lights, a red and a green on each side, one and one-half inches apart with 12 inches between each pair of lights. One inch below each pair of lights was a large black button centered between the lights. Either button, when pressed, would terminate power to any and all lights on the board and the electrical interval timer. A digital second timer recorded resolution times to hundredths of a second, and the timer was located out of *S*'s view throughout the experiment. A third button, centered approximately one inch from the edge of the board in front of *S*, served as a starting point for *S*'s finger. Located on *E*'s side of the center partition were switches, which enabled *E* to control the various combinations of red and green lights on *S*'s side of the board and to activate the timer.



### 3. Procedure

In order to assure that no *E* bias occurred in data collection, the *Ss* from each of the two drive groups were scheduled to appear randomly, and *E* was unaware of the group to which *S* belonged until data collection was completed.

The following instructions were given *S*: "In front of you is a board with four lights on it. As you can see there are a red and a green light on each side. When I signal you by saying OK, you will press the button directly in front of you. Please use only the forefinger of one hand and keep your nonpreferred hand in your lap. A short time after you have pressed the button in front of you, one or more of the lights on the board will come on. If a green light comes on, you are to trace along the line on the board with your forefinger to the button below that light and press that button. If a red light comes on, you are to trace with your forefinger along the line to the button on the side opposite the red light and press the button there. In other words, you are to trace a line toward a green light should it come on and away from a red light should it come on. It is very important that your forefinger remain on the small black start button in front of you until you are absolutely sure of where you plan to trace on the board. The amount of time between when your finger leaves the start button and when it reaches one of the large black destination buttons should be kept at the very minimum. To do your best think about exactly where you intend to trace with your forefinger before it leaves the start button. This is very important. Now, are there any questions before we begin? I can answer no questions once we have started. OK, we are ready; so please press the start button and we will begin."

Counterbalancing of the three conflicts was accomplished with a complete between-subject model, and each conflict type occurred equally often at each stage of practice and preceded and followed the other conflict types an equal number of times. The experimental design was a  $2 \times 3$  with repeated measures on the second factor (conflict type), and there were 20 *Ss* in each of the two levels of the first factor (manifest anxiety).

For the AP-AP conflict *S* received a series of 10 practice trials in which he randomly received a single green light on either the left side of the board or the right side. The digital timer was activated on each practice trial although no record was kept of practice trial times. There was a five-second intertrial interval during which *S* was asked to repress the start button in

front of him. On trial 11 both green lights were activated, thus representing an AP-AP conflict. Conflict resolution time, consisting of the time interval between the activation of the two green lights and *S*'s depression of one of the buttons below the lights, was recorded to hundredths of a second. The same procedure was followed for the AV-AV conflict as for the AP-AP conflict, except that a single red light was activated on either side of the board on the 10 practice trials with both red lights on trial 11 representing the AV-AV conflict. Again, conflict resolution time on trial 11 was recorded.

The procedure for the DAP-AV conflict was identical to the two above types except that there was a series of 20 practice trials involving either a left green, right red, right green, or left red light which preceded trial 21 when all four lights appeared simultaneously. As before, resolution time was recorded on the test trial. The additional practice trials for the DAP-AV conflict followed Hovland and Sears' (2) recommendation that because of the alleged degree of difficulty of the DAP-AV conflict *S* should be given double the number of practice trials prior to presentation of the DAP-AV conflict.

### C. RESULTS

An analysis of variance of the resolution times presented in Figure 1 yielded a nonsignificant Conflict  $\times$  MAS interaction. However, significant main effects for both the MAS factor,  $F(1, 38) = 24.30, p < .05$ , and the Conflict factor  $F(2, 76) = 14.54, p < .05$ , were obtained. A significant main effect for the MAS factor indicated, as hypothesized, that the HD group took significantly longer to resolve the conflict types than did the LD group (2.24 seconds compared to 1.15 seconds).

The Duncan test for differences among ordered means at the .05 level was performed on the significant Conflict factor. It indicated that the AP-AP (1.32 seconds) conflict was resolved more rapidly than either AV-AV (1.77 seconds) or DAP-AV (2.00 seconds) conflicts, but that the two later conflicts did not differ significantly.

### D. DISCUSSION

Since the present study was not concerned with the various modes of conflict resolution but only with speed of resolution, it was difficult to make a direct comparison with Hovland and Sears (2). One of their findings was confirmed; namely, AP-AP conflict was easier to resolve than either AV-AV or DAP-AV conflicts. On the other hand, the present finding of no difference in resolution time between AV-AV and DAP-AV was at variance

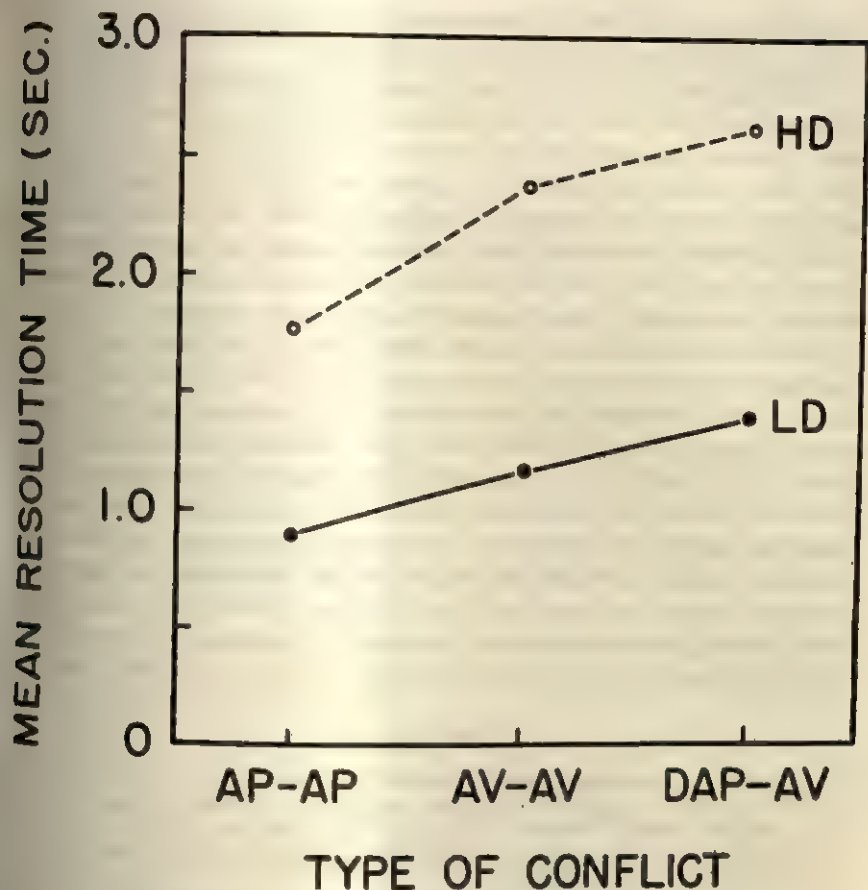


FIGURE 1  
MEAN CONFLICT RESOLUTION TIME FOR HIGH DRIVE (HD) AND LOW DRIVE (LD)  
SUBJECTS FOR APPROACH-APPROACH (AP-AP), AVOIDANCE-AVOIDANCE (AV-AV),  
AND DOUBLE APPROACH-AVOIDANCE (DAP-AV) MOTOR CONFLICTS

with the conclusion of Hovland and Sears that the DAP-AV conflict was the most difficult conflict to resolve.

The HD and LD comparisons yielded data consistent with past research (3, 11) in that manifest anxiety was inversely related to performance in complex conflict situations. Taylor (9) indicated that the differential performance of HD and LD groups in a relatively complex task was dependent upon the number and comparative strengths of the competing response

tendencies, and apparently the motor conflict situation was sufficiently complex to yield differences in the anxiety conditions.

Further support was given by the present study to Taylor's (10) theoretical statement that MAS scores reflect differences in a chronic emotional state. Whatever the activities or processes involved in motor conflict resolution, they are strongly influenced by the individual's pre-existing level of generalized drive. The theoretical implication seems to be that regardless of the undefined cognitive or motor processes involved in conflict resolution, the fact that an individual has a relatively high level of generalized drive will increase the amount of time required to solve motor conflicts.

Several areas of follow-up research were indicated, and they included (a) investigation of personality characteristics of the "kinds of persons who respond with indecision and uncertainty under minimal conflict or with speed, dispatch, and lack of vacillation under conflicts of considerable complexity" (4, p. 489); (b) examination of the effects of modeling on the behavior of HD and LD Ss in conflict resolution by means of a verbal conflict paradigm reported by Powell (7); (c) numerous situational variables, such as fear of failure, threat in the form of noxious stimulation, or motivating instructions might be considered; and (d) future experiments should consider blending both qualitative and quantitative aspects and looking at correlations between the two response measures for the major types of conflict.

#### REFERENCES

1. CHILD, I. L. Personality. *Ann. Rev. Psychol.*, 1954, 5, 149-170.
2. HOVLAND, C. I., & SEARS, R. R. Experiments on motor conflict: I. Types of conflict and their modes of resolution. *J. Exper. Psychol.*, 1938, 23, 477-493.
3. KERRICK, J. S. Some correlations of the Taylor Manifest Anxiety Scale. *J. Abn. & Soc. Psychol.*, 1955, 50, 75-77.
4. KIMBLE, G. A., & GARMEZY, N. *Principles of General Psychology* (2nd ed.). New York: Ronald Press, 1963.
5. LEWIN, K. Environmental forces in child behavior and development. In C. Murchison (Ed.), *A Handbook of Child Psychology*. Worcester, Mass.: Clark Univ. Press, 1931.
6. ———. *A Dynamic Theory of Personality*. New York: McGraw-Hill, 1935.
7. POWELL, N. R. Influence of vicarious reinforcement on an observer's speed of conflict resolution. Unpublished Master's thesis, University of Richmond, Richmond, Virginia, 1971.
8. TAYLOR, J. A. The relationship of anxiety to the conditioned eyelid response. *J. Exper. Psychol.*, 1951, 41, 81-92.
9. ———. A personality scale of manifest anxiety. *J. Abn. & Soc. Psychol.*, 1953, 48, 285-290.
10. ———. Drive theory and manifest anxiety. *Psychol. Bull.*, 1956, 53, 303-320.

11. TAYLOR, J. A., & SPENCE, K. W. The relationship of anxiety level to performance in serial learning. *J. Exper. Psychol.*, 1952, **44**, 61-64.
12. WENAR, C. Reaction time as a function of manifest anxiety and stimulus intensity. *J. Abn. & Soc. Psychol.*, 1954, **49**, 335-340.

*Department of Psychology*  
*University of Richmond*  
*Richmond, Virginia 23173*





## CONSERVATIVE ATTITUDES AND AUTHORITARIAN VALUES\*<sup>1</sup>

*Catholic University*

---

ANTANAS SUZIEDELIS AND MAURICE LORR

---

### SUMMARY

The aim of the study was to determine the structure and dimensionality of social-political attitudes. Also investigated was the hypothesis that the liberal *vs.* conservative attitude dimension was synonymous with the value dimension rule free *vs.* authoritarian. An attitude measure consisting of simply stated social-political issues and a personal values inventory were administered to a varied sample of 300 men and women. Analysis of the attitude items disclosed three factors—Liberal *vs.* Conservative, Freedom of Sex Expression, and Equalitarianism—whose intercorrelations indicated a common dimension of conservatism. The cross-correlations in both the male and female samples support the hypothesis that conservative attitudes and authoritarian values were essentially identical constructs. Groups ranked by conservatism and groups ranked as to age increased significantly in mean conservatism score as anticipated.

---

### A. INTRODUCTION

The purpose of the study was to investigate the structure and dimensionality of the domain of social-political attitudes in the United States. Two hypotheses were examined within the context of the study. The first hypothesis was that one inclusive factor of liberalism-conservatism could account for the relation among the factors isolated. The second hypothesis was that the personal value dimension, rule free *vs.* authoritarian, was essentially the same as the liberal-conservative continuum.

There are at least two components involved in attitude measurement: One is affective and the other is cognitive. An individual learns to believe that a given behavior, policy, or group is characterized by certain outcomes. This belief regarding the characteristics of the referent can be measured by asking the subject to indicate his degree of agreement or disagreement with the issue.

---

\* Received in the Editorial Office on January 16, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Received partial support from NIH grant MH-17571.

The individual also learns to respond with positive or negative affect toward the referent. The affect can be assessed by asking the individual how favorably or unfavorably he is disposed toward the issue. An attitude may be defined as the predisposition to respond favorably or unfavorably to a referent.

An important problem is how to measure attitudes. With but few exceptions most attitude scales are defined by complex statements somewhat ambiguous in meaning. As in the case of the much investigated F Scale, the acquiescent response is likely to be enhanced. An alternative approach is to use one or two words to indicate each referent. Wilson and Patterson (9) used such categories as Socialism, Church Authority, and Censorship to characterize controversial issues. The approach taken here was similar except that the issues were described with greater explicitness by means of short phrases. The subject was asked whether he favored or opposed such matters as tax on church property, complete freedom of speech, civil disobedience as a form of protest.

## B. METHOD

### 1. *The Attitude Form*

The attitude form was composed of 115 items describing issues that the subject was to indicate whether he "favored" or "opposed." A "?" answer was also included for the respondent who had no opinion or who lacked information to respond. Examples of the issues may be found in Tables 1 and 2.

Ten political and social categories were constructed following a survey of problems currently being debated. Approximately 10 or more items were written to represent each bipolar continuum proposed. The postulated belief categories were as follows: Religious *vs.* Antireligious, Promilitary *vs.* Pacifist, Equalitarian *vs.* Elitist, Private *vs.* Public Ownership of Resources, Sex Freedom *vs.* Conformity, Freedom of Expression *vs.* Restraint, Nationalist *vs.* Internationalist, Individual Freedom *vs.* Law and Order. In addition, items representing unipolar categories of Pro-Public Welfare and Environmental Control were included.

### 2. *Inventory of Personal Values*

Members of the sample studied also completed a 12-dimensional inventory<sup>2</sup> of personal values, as well as the attitude form (see list of value scales on Table 3). One aim was to test the hypothesis that the liberal-conservative

<sup>2</sup> Unpublished Personal Value Inventory developed by M. Lorr and A. Suziedelis.

continuum was closely associated with the Authoritarian value conception. The referent values consisted of statements descriptive of a wide range of end-states or means-to-ends judged important or unimportant to the individual. Such personal values serve as criteria or standards that guide judgment and action. The social attitudes, on the other hand, are less general in their influence and are more closely tied to knowledge about specific issues.

### 3. Subjects

The subjects consisted of 300 persons (158 men and 142 women) from a wide range of sources. Included were such groups as firemen, enlisted soldiers, college students, members of a Quaker group, nuns, defense engineers, actors, members of an old age home, a hippie group, a sample of Mormons, and nurses. The data were collected in the spring of 1972.

### 4. Procedure

The 115 items were intercorrelated and analyzed by the method of principal components with units on the diagonals. Cattell's (1) "scree" test was used to determine the number of reliable factors present. Kaiser's normal varimax procedure was then applied to the 12 factors retained.

The next step consisted of grouping items that correlated at least .40 with one of the factors into 20 subtests of three to five items. All subjects were then scored on the 20 subtests, each score consisting of the sum of unit-weighted items. The correlations among the 20 variables with units in the matrix diagonals were again factored. Inspection of the latent roots revealed a sharp drop after three factors that accounted for 57% of the variance. The varimax rotational solution obtained made it evident that a correlated solution would provide a better fit to the data. Accordingly an oblique promax solution (7) was obtained.

The last step preceding assignment of items to the three attitude scales consisted in correlating each item with the three total scores. Items that correlated at least .40 with one of the scales and at least .10 lower with the other two measures were added to those already included. The final key consisted of 60 items, of which 40 defined the Liberal-Conservative measure, 12 the Sex Freedom measure, and eight the Equalitarian scale. For the men the Kuder-Richardson reliabilities were .96, .90, and .77, while for the women these were .94, .91, and .81, respectively.

## C. RESULTS

### 1. The Attitude Scales

The dominating factor is defined by 32 items that correlate with it  $\pm .35$  or more. The final scale of 40 items included those added on the basis of cor-

**TABLE 1**  
**REFERENT ISSUES THAT DEFINE THE LIBERAL-CONSERVATIVE DIMENSION**

Liberal	Conservative
Prohibition of all nuclear arms Public ownership of all large corporations Socialized medicine Reduction of military spending Civil disobedience as a form of protest Disarmament treaty with Russia Curbing police powers A socialistic society Eliminate use of Bible at swearing-in ceremonies Conscientious objector status on non-religious grounds Busing to achieve integration in schools Black Panthers Amnesty (no penalty) for draft evaders Limitations on FBI activities Limits on stop and frisk procedure Elimination of the draft Reduction of CIA activities More control of wiretapping Take out "In God We Trust" from coins/currency Sex education in elementary schools	Development of ABM Use of wiretapping by police against criminals Maintain strong armed forces Military intervention in other countries where our national interest is at stake Right of business to set dress regulations University research on defense contracts U.S. military bases in foreign countries Loyalty tests for some jobs More police on patrol Prosecution of persons revealing government secrets No-knock laws to allow seizure of criminal evidence ROTC in colleges and universities Regulations of press to safeguard national interests Screening of radicals from government positions Death penalty Military foreign aid Screening of Communists from government positions Separate system of military justice Prayers in public schools Strong penalties for civil disobedience

relation with the total score. Inspection of the items in Table 1 indicates that the range of issues is very broad. Included are concerns about law and order, the military, civil liberties, religion, and public welfare. Thus it is plausible to hypothesize the factor to represent a dimension of liberalism *vs.* conservatism. As it may be seen, the scale is balanced for the number of liberal and conservative referents.

The second factor, shown on Table 2, is primarily concerned with freedom of sexual expression. If items descriptive of more conventional restraint and conformity were added, the dimension would undoubtedly become bipolar. The third and smallest dimension, also shown on Table 2, is concerned with equal rights for both race and women. In this respect the factor is very similar or equivalent to Comrey and Newmeyer's (2) dimension of Racial Tolerance. Their factor is defined both by equal rights items (called Social Change) and by racial tolerance items.

The correlations among the three scales scores are presented in Table 3



TABLE 2  
REFERENT ISSUES THAT DEFINE FREEDOM OF SEX EXPRESSION AND EQUALITARIANISM

Freedom of Sex Expression	Equalitarianism
Legalized prostitution	Adopting children of a different race
Living together without marriage	No restrictions on mixed marriages
Birth control clinics for minors	Full equality for women
Communal living	Women judges
Extramartial sex	Segregated neighborhoods
Contract marriage (limited duration)	Interracial marriages
No censorship of pornography	Open housing (no restrictions on sales or rentals)
Legalized abortion	Opening men-only jobs to women
Easier divorce laws	
Closing down strip shows and smut shops	
Premarital sex	
Coeducational dorms	

for both men and women. The correlations among the primary factors themselves, obtained in the promax analysis, were slightly lower ( $r_{12} = .55$ ,  $r_{13} = .56$ ,  $r_{23} = .45$ ). The results suggest that a broad higher level factor probably underlies the three factors. Comrey and Newmeyer (2) also found only one major second-order factor when they analyzed the correlations among their nine primary factors.

## 2. Scale Validity

One approach to evaluating scale variables is to compare groups of known position with regard to their mean scores. Fourteen groups judged to occupy widely varying positions on the liberal-conservative continuum were selected for comparison. Table 4 lists these groups ( $N = 118$ ) by mean score. The eta coefficient of .84 indicates that the groups were sharply differentiated in liberalism-conservatism. The Quaker church group, a gay lib group, and a hippie group identified the liberal end. An East European ethnic group, a company of firemen in a firehouse, and a Lutheran fundamentalist group defined the conservative end of the continuum. The groups were also differentiated at significant  $F$  levels by Freedom of Sex Expression (15.4) and by Equalitarianism (7.3).

The liberal-conservative variable is known to be associated with age. Older people tend to be more conservative than younger people. Two samples of 120 cases were constructed to be matched roughly for age, education, and source. The samples were then divided into six age groups extending from 15 years to over 60. There was a highly significant relation between liberalism-conservatism and age in women, but in men it reached only marginal significance ( $p < .056$ ). However, it should be noted that the

TABLE 3  
CROSS-CORRELATIONS OF ATTITUDE MEASURES WITH 12 VALUE SCALES  
(DECIMALS OMITTED)

Variable	Men			Women		
	Lib	Sex	Equal	Lib	Sex	Equal
Value scales						
Socially Concerned	40	04	36	-01	-07	21
Intellectual	12	-05	19	-07	03	03
Hedonistic	10	33	00	05	16	05
Self Interested	-09	13	-15	01	14	-02
Authoritarian	-79	-55	-55	-77	-64	-48
Religious	-42	-62	-37	-60	-76	-39
Elitist	-40	-37	-33	-43	-39	-43
Work Ethic	-32	-18	-22	-35	-29	-35
Independent	-07	07	-04	-04	15	-05
Stoic	-30	-09	-20	-32	-16	-24
Venturesome	01	22	09	10	34	10
Conscientious	-04	-12	-01	-25	-12	03
Attitude measures						
Lib	1.00	61	69	1.00	63	49
Sex	61	1.00	54	63	1.00	45
Equal	69	54	1.00	49	45	1.00

Note: Lib = Liberal-Conservative measure; Sex = Freedom of Sex Expression measure; Equal = Equalitarian measure.

samples were by no means representative of age groups in the population as a whole. One would expect liberalism-conservatism would decrease monotonically with age. It did not decrease uniformly in the groups represented here.

TABLE 4  
MEAN LIBERAL-CONSERVATIVE SCORES OF 14 SMALL GROUPS

Group	N	Mean score
Quakers	22	27.5
Gay lib	9	23.3
Hippies	10	23.1
Chess club	12	15.7
University students	24	15.5
City college students	15	8.4
Suburbanites	10	-2.3
Seventh Day Adventists	11	-6.4
Old people	11	-12.0
WACs	7	-13.1
Defense engineers	11	-17.4
Lutherans	12	-17.4
Firemen	20	-18.9
East European ethnics	14	-21.9

Note:  $F$  ratio = 32.3,  $\eta^2$  = .84,  $df$  = 13/174.

### 3. Relations to Value Dimensions

Table 3 presents the cross-correlations of each of the 12 value measures with the attitude variables. The correlations clearly support the conjecture that the liberal-conservative attitude dimension is very similar to, or identical with, the Authoritarian value dimension in both men and women ( $-.79$  and  $-.77$ ). Fundamentalist Religious belief and Stoic values are also strongly associated with liberalism-conservatism. Adherence to the Work Ethic and Elitist viewpoints are somewhat less associated with liberalism-conservatism. The other two measures of attitude display similar correlational patterns. The scale intercorrelations of the three attitude measures indicate the presence of a common factor in both men and women. Within the table a correlation of  $.21$  is significant at the  $.01$  level.

### D. DISCUSSION

The results indicate that the use of simple referents to define social issues is relatively effective. However, there remains the question regarding the similarity of findings to prior analyses. Ferguson (5) isolated three social attitudes: Religionism, Humanitarianism, and Nationalist. Eysenck (4) found a Radicalism-Conservatism factor R and a Tenderminded-Toughminded factor T. However, Defronzo (3) demonstrated that the T factor can be resolved into one concerned with Religion and Humanitarianism. Kerlinger (8), in a very recent study that just came to our attention, found three liberal and three conservative factors. The conservative dimensions were Religiosity, Economic Conservatism, and Educational Traditionalism. The liberal dimensions were called Civil Rights, Social Liberalism, and Child Centered Education. Two second-order factors, rather than one, were demonstrated in keeping with his hypothesis that social attitudes were dualistic.

It is very possible that the type of items used will influence the dimensionality of responses obtained. As Fishbein and Ajzen (6) indicate, it makes a difference as to whether the subjects are asked their degree of agreement or their degree of favorability toward the issues presented. The complexity of the statement (the cognitive element) obviously plays a similar role. Investigations seem needed to ascertain the roles of these factors.

### REFERENCES

- CATTELL, R. B. The scree test for the number of factors. *Multivar. Behav. Res.*, 1966, 1, 245-276.
- COMREY, A. L., & NEWMAYER, J. A. Measurement of radicalism-conservatism. *J. Soc. Psychol.*, 1965, 67, 357-369.

3. DEFRONZO, J. Religion and humanitarianism in Eysenck's T-Dimension and left-right political orientation. *J. Personal. & Soc. Psychol.*, 1972, **21**, 265-269.
4. EYSENCK, H. J. *The Psychology of Politics*. London: Routledge & Kegan Paul, 1954.
5. FERGUSON, L. W. The measurement of primary social attitudes. *J. of Psychol.*, 1940, **10**, 199-205.
6. FISHBEIN, M., & AJZEN, I. Attitudes and opinions. *Ann. Rev. Psychol.*, 1972, **23**, 487-544.
7. HENDRICKSON, A. E., & WHITE, P. O. Promax: A quick method for rotation to oblique structure. *Brit. J. Stat. Psychol.*, 1964, **17**, 65-70.
8. KERLINGER, F. N. The structure and content of social attitude referents: A preliminary study. *Educ. & Psychol. Meas.*, 1972, **32**, 613-630.
9. WILSON, G. D., & PATTERSON, J. R. A new measure of conservatism. *Brit. J. Soc. & Clin. Psychol.*, 1968, **7**, 264-269.

*Department of Psychology*  
*Catholic University*  
*Washington, D. C. 20017*

## GROUP DESENSITIZATION OF TEST ANXIETY IN ELEMENTARY SCHOOL\*<sup>1</sup>

*State University College at Buffalo*

---

ARREED F. BARABASZ

---

### SUMMARY

Fifth and sixth grade students ( $N = 87$ ) were divided into High Test Anxious and Low Test Anxious groups on the basis of galvanic skin responses (GSR). Ss identified as High Anxious responded positively to group systematic desensitization procedures, exhibiting significantly ( $p < .05$ ) lower autonomic indices of test anxiety and significantly ( $p < .001$ ) improved criterion test scores. Low Anxious Ss exposed to the systematic desensitization program demonstrated no significant change in criterion test scores.

---

### A. INTRODUCTION

Several investigations have shown the debilitating effects of test anxiety on test scores (6, 14, 17, 18). This negative relationship exists at all elementary school levels (17, 19). Highly anxious students receive lower grades and have a higher academic failure rate than nonanxious or low anxious students of equivalent intelligence (20, 21). Paul and Eriksen (15) support the notion that high test anxious students suffer impaired performance under regular examination or testing conditions because of heightened physiological activity and self-depreciating ruminations. Negroes have been shown to be significantly more test anxious than Caucasians on the basis of galvanic skin response (GSR) comparisons (1).

Wolpe (24) described a technique of psychotherapy by reciprocal inhibition in which anxiety responses are extinguished when aroused in temporal contiguity with a psychological state antagonistic to anxiety. This procedure of systematic desensitization consists of gradual introduction of progressively stronger phobic stimuli while deep relaxation is employed as the anxiety inhibiting state. Autonomic effects accompanying deep relaxation have been

---

\* Received in the Editorial Office on January 19, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The project was supported by a grant from the Research Foundation of the State University of New York, Contract No. 021-7126-A.



demonstrated to be diametrically opposed to those characteristic of anxiety. Jacobson (7, 8) discovered a decrease in pulse rate and blood pressure coincident with deep muscle relaxation. Galvanic skin resistance increases while respiration becomes slower and more regular during relaxation (3, 25). Systematic desensitization has been demonstrated to be highly effective in controlled experimental studies, as well as in clinical experience with adults individually or in groups or with children individually (2, 9, 10, 11, 12, 16, 24, 25).

Few studies have dealt with anxiety related to school. Paul and Shannon (16) desensitized college males by group method and found a significant improvement in grade point average. Only one study has been located in the literature which deals directly with group desensitization of test anxiety. Cohen (4) desensitized undergraduates who had scored in the upper 30 percent on the Test Anxiety Scale. Results indicated a significant decrease in test anxiety scores and a significant increase in grade point average for the experimental subjects. Children have also been successfully treated with similar techniques in individual settings (12), but group desensitization of test anxiety has not been attempted.

The purpose of this investigation was to determine whether or not fifth and sixth grade elementary school students could be successfully counterconditioned to test anxiety in their regular homeroom settings with a resultant reduction in physiologically measured anxiety as well as raised criterion test scores. On the basis of Wolpe's theory of reciprocal inhibition it was hypothesized that test anxious subjects exposed to the systematic desensitization program would exhibit lower autonomic indices of test anxiety and show a significant improvement of scores on the criterion test.

## B. METHOD

### 1. Subjects

Ss consisted of fifth and sixth grade students enrolled at the Campus School, State University College at Buffalo, New York ( $N = 87$ ). The sample represented both middle and lower socioeconomic levels, as well as both Negroes and Caucasians, as determined by occupations of Ss' fathers (22). Ss were randomly assigned to homeroom groups at the beginning of the school year.

### 2. Procedure

To provide a pretreatment criterion measure the Lorge-Thorndike Intelligence Test was simultaneously administered to all groups of students in their

regular classrooms via video taped instructions. Ss were told that the test results would be placed on their permanent record cards. [Mild threat in academic and test situations serves to elicit anxiety (5, 13, 23).]

Ss were divided into High Test Anxious and Low Test Anxious groups with use of the Lafayette 76012 Polygraph Model C. Galvanic skin resistance was transmitted via finger electrodes, and heart rate via photoelectric finger plethysmograph. Electrodes were attached to the volar surface of the proximal phalanx on the third and fourth digits on each of the S's hands. Ss were exposed to an audio tape recording which asked them to visualize the following 15 situations. (There was a 10 second pause between each stimulus. The stimulus designed to arouse test anxiety was placed last so as to provide ample time for establishing each S's autonomic response baseline.)

1. Picture yourself sitting in front of a fireplace on a cold winter's day.
2. Picture yourself by the ocean on a summer day with a warm breeze blowing through you hair.
3. Picture yourself in your favorite chair watching television.
4. Picture yourself resting on the couch after a large meal.
5. Picture yourself lying on the grass fishing in a still pond.
6. Picture yourself lying under a tree listening to a little brook.
7. Picture yourself in front of a fish tank watching the fish slowly swim by.
8. Picture yourself lying on the beach watching the sea gulls glide slowly in the sky.
9. Picture yourself watching large snow flakes drift to the ground.
10. Picture yourself watching a kite fly in the sky.
11. Picture yourself lying on a rug listening to the radio.
12. Picture yourself watching two rabbits play in the back yard.
13. Picture yourself lying in bed and looking out your window watching a robin build a nest.
14. Picture yourself lying in a leaf pile watching the leaves spiral down on an autumn day.
15. Picture yourself taking a difficult examination that determines whether you pass or fail.

The experimental groups consisted of one sixth grade and one fifth grade comprised of High Test Anxious and Low Test Anxious Ss. The groups were exposed to a systematic desensitization program on five consecutive days. Ss participated in their regular homeroom settings with the instructions of E and with a reduction in room lighting. Wolpe's (26) relaxation training procedure was employed in conjunction with a test anxiety hierarchy constructed by the

investigator. The first two days stressed only relaxation training. The test anxiety hierarchy was gradually introduced during the remaining three days. *E*'s speed of progression through the list was modified by *S*'s input. *Ss* indicated feelings of anxiety in response to a stimulus on the hierarchy by raising their hands. When stimuli evoked more than two hands raised, *E* repeated the prior stimulus in the hierarchy emphasizing relaxation. Brief repetitions of relaxation inducing stimuli were reintroduced until the two hand maximum was attained.<sup>2</sup>

The control group consisted of one sixth grade and one fifth grade comprised of High Test Anxious and Low Test Anxious *Ss*. This group was not exposed to the desensitization program.

Upon termination of treatments the Lorge-Thorndike Intelligence Test (Form 2) was administered to experimental and control groups. Once again *Ss* were told that the results would be placed on their permanent record card. All *Ss* were then re-examined by the polygraph as outlined previously.

### C. RESULTS

In order to determine whether or not the systematic desensitization program influenced *Ss*' galvanic skin resistance, *t* tests were calculated comparing the posttreatment scores of Experimental High Anxious *versus* Control High Anxious and Experimental Low Anxious *versus* Control Low Anxious *Ss*. The results appear in Table 1.

The results presented in Table 1 demonstrate that Experimental *Ss* pretreatment classified as High Anxious had significantly ( $p < .05$ ) lower GSR scores than the High Anxious Controls. Computation of omega square showed 14% estimate of the degree of statistical association. The GSR scores of Experimental and Control *Ss* pretreatment classified as Low Anxious were not found to differ significantly.

Heart rate data failed to differentiate between High and Low Anxious *Ss* on the pretreatment polygraph tests. Analysis for heart rate was, therefore, not attempted.

In order to determine whether or not the desensitization program resulted in an improvement in the Lorge-Thorndike criterion test scores, analysis of variance (Kruskal-Wallis) was computed on the raw equivalency change scores for Experimental High Anxious, Experimental Low Anxious, Control High Anxious, and Control Low Anxious *Ss*. The Lorge-Thorndike equivalency table equated scores for the parallel forms administered. Change in scores was determined on a pre-post test basis for each *S*. Nonnormal Lorge-Thorndike

<sup>2</sup> Anxiety hierarchy transcripts may be obtained by writing the author at the address shown at the end of this article.

TABLE 1  
CONTROL AND EXPERIMENTAL POSTTREATMENT GALVANIC SKIN RESPONSE COMPARISONS  
FOR HIGH AND LOW ANXIOUS SUBJECTS

Group	<i>N</i>	$\bar{X}$	<i>t</i>	Omega square
Experimental High Anxious	20	1.28	2.64*	.14
vs. Control High Anxious	17	2.79		
Experimental Low Anxious	22	1.26	.25	
vs. Control Low Anxious	24	1.35		

\*  $p < .05$ .

score distributions, due to the high-low anxiety blocking, and unequal variances necessitated the use of the nonparametric test. The Kruskal-Wallis analysis of variance yielded a significant  $H$  of 16.64 ( $p < .001$ ) for the overall comparison. Experimental Ss pretreatment classified as High Anxious showed a significant improvement in criterion scores as contrasted with Control High Anxious Ss ( $H = 9.39$ ,  $p < .01$ ). No significant difference in criterion scores was demonstrated for the Experimental Low Anxious Ss as compared with the Control Low Anxious Ss ( $H = .34$ ,  $p > .05$ ).

#### D. DISCUSSION

The results support the hypothesis that test anxious Ss exposed to the systematic desensitization program would exhibit lower autonomic indices of test anxiety and show a significant improvement of scores on the criterion measure. These findings are consistent with Wolpe's theory of reciprocal inhibition. High Test Anxious Ss exposed to the systematic desensitization program exhibited significantly lower anxiety scores, as measured by galvanic skin resistance, than the High Test Anxious Controls, while no significant difference was found between Low Anxious Experimentals and Low Anxious Controls. High Test Anxious Ss exposed to the desensitization program demonstrated a significant improvement in their Lorge-Thorndike criterion test scores, while no significant changes resulted for the other groups of Ss. It is of particular interest that the criterion test scores of the Low Anxious Experimental Ss did not change significantly. Evidently the relaxation training did not impair these Ss' test taking motivation. Implementation of classroom programs for test anxiety desensitization is apparently simplified, since there seems to be no indication that Low Test Anxious Ss would have to be segre-



gated from the treatment settings. Further replication is, however, recommended. Differential spacing or massing of desensitization sessions and the feasibility of training classroom teachers to carry out the program might also be studied.

### REFERENCES

1. BARABASZ, A. F. Galvanic skin response and test anxiety among Negroes and Caucasians. *Child Stud. J.*, 1970, 1, 33.
2. BURNETT, A., & RYAN, E. Conditioning techniques in psychotherapy. *Can. Psychiat. Assoc. J.*, 1964, 9, 140.
3. CLARK, D. F. The treatment of monosymptomatic phobia by systematic desensitization. *Behav. Res. Ther.*, 1963, 1, 63.
4. COHEN, R. Group desensitization of test anxiety. Unpublished Ph.D. thesis, Syracuse University, Syracuse, New York, 1968.
5. DIAMOND, L. Effects of anxiety and stress on children's performance. Unpublished Doctoral dissertation, Catholic University of America, Washington, D. C., 1965.
6. HILL, K. T., & SARASON, S. B. The relation of test anxiety and defensiveness to test and school performance over the elementary school years: A further longitudinal study. *Monog. Soc. Res. Child Devel.*, 1966, 31, Whole No. 104.
7. JACOBSON, E. Variation of blood pressure with skeletal muscle tension and relaxation. *Amer. Int. Med.*, 1939, 12, 1194.
8. ———. Variation of pulse rate with skeletal muscle tension and relaxation. *Amer. Int. Med.*, 1940, 13, 1619.
9. LANG, P. J., LAZOVIC, A. D., & REYNOLDS, D. Desensitization, suggestibility and pseudotherapy. *J. Abn. Psychol.*, 1965, 70, 395.
10. LAZARUS, A. A. Group therapy of phobic disorders by systematic desensitization. *J. Abn. & Soc. Psychol.*, 1961, 63, 504.
11. ———. The results of behavior therapy in 126 cases of severe neurosis. *Behav. Res. Ther.*, 1963, 1, 1969.
12. LAZARUS, A. A., & ABRAMOVITZ, A. The use of emotive imagery in the treatment of children's phobia. *J. Ment. Sci.*, 1962, 108, 191.
13. MEISELS, M. Test anxiety, stress, and verbal behavior. *J. Consult. Psychol.*, 1961, 31, 577.
14. O'REILLY, R. Conceptualizing questionnaire research on anxiety in the school culture: A theoretical and methodological viewpoint. Unpublished Master's thesis, Cornell University, Ithaca, New York, 1966.
15. PAUL, G. L., & ERIKSEN, C. W. Effects of test anxiety on real life examination. *J. Personal.*, 1964, 32, 480-494.
16. PAUL, G. L., & SHANNON, D. T. Treatment of anxiety through systematic desensitization in therapy groups. *J. Abn. Psychol.*, 1966, 71, 124.
17. RUEBUSH, B. K. Anxiety. In H. W. Stevenson (Ed.), *Child Psychology: The Sixty-Second Yearbook of the National Society for the Study of Education*. Chicago: Univ. Chicago Press, 1963.
18. SARASON, S. B., HILL, T. K., & ZIMBARDO, P. G. A longitudinal study of the relation of test anxiety to performance on intelligence and achievement tests. *Monog. Soc. Res. Child Devel.*, 1964, 29, 4-49.
19. SARASON, S. B., et al. Anxiety in elementary school children: A report of research. New York: Wiley, 1960.
20. SPIELBERGER, C. D. The effects of manifest anxiety on the academic achievement of college students. *Ment. Hyg.*, 1962, 46, 420-426.



21. SPIELBERGER, C. D., & KATZMEYER, W. G. Manifest anxiety, intelligence, and college grades. *J. Consult. Psychol.*, 1959, **23**, 278.
22. WARNER, W., MEEKER, M., EELLS, K. Social Class in America. Chicago: Sci. Res. Assoc., 1949.
23. WINKEL, G. H., & SARASON, E. G. Subject, experimenter, and situational variables in research on anxiety. *J. Abn. & Soc. Psychol.*, 1964, **68**, 601.
24. WOLPE, J. Psychotherapy by Reciprocal Inhibition. Stanford, Calif.: Stanford Univ. Press, 1958.
25. ———. Behavior therapy in complex neurotic states. *Brit. J. Psychiat.*, 1964, **110**, 28.
26. WOLPE, J., & LAZARUS, A. A. Behavior Therapy Techniques. New York: Pergamon Press, 1966.

*Child Study Center*  
*State University College at Buffalo*  
*1300 Elmwood Avenue*  
*Buffalo, New York 14222*



## TRAINING OF NUMBER CONSERVATION IN RETARDATE<sup>\*</sup> 1,2

*Eastern Washington State College*

---

ROY P. LANCASTER AND DONALD L. McMANIS

---

### SUMMARY

Eighteen retarded adults, divided into three groups equated for *IQ*, *MA*, and number conservation performance, received addition-subtraction and reversibility training under either cognitive-conflict or nonconflict conditions, or received no training. Under cognitive-conflict, transformation of one of two sets of discrete elements produced a perceptual illusion, and the addition-subtraction and reversibility operations applied to that set produced conflict between its length and density. Under nonconflict, simultaneous application of the operations to both sets avoided such conflict. Both training groups made significant gains from pretest and significantly exceeded controls in posttest number conservation ( $ps < .05$ ). Lack of differential gains by the training groups suggested that cognitive conflict is not essential to induce number conservation.

---

### A. INTRODUCTION

According to Piaget (4), the development of number conservation is one of the major accomplishments demarking the transition from preoperational to concrete operational thinking in normal children, occurring at about seven years of age. Prior to this age, children believe that the number of elements in a set of objects can vary if the spatial distribution of the elements is modified. At about age seven, the child acquires the notion of invariance of number in the face of illusory perceptual cues, with this notion being based on logical necessity. Piaget asserts that the idea of invariance of number arises as a result of the internal process of equilibration, which involves the reciprocal subordinate processes of assimilation and accommodation. As the child's interactions with his environment confront him with problems whose solutions

---

\* Received in the Editorial Office on January 19, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Preparation of the manuscript for the present study was supported by the Human Development Center, Eastern Washington State College.

<sup>2</sup> Direct all correspondence regarding the present study to the second author at the address shown at the end of this article.

require cognitive schemas which are slightly more advanced than his current development will permit, but do not require major modifications of existing operations, his current schemas will incorporate (assimilate) the necessary information from the environment. As assimilation takes place, the schemas will become modified (accommodation) sufficiently to permit solution of problems previously beyond the child's capacity.

This point of view differs from that of learning theory, since the beneficial effects of experience in interacting with the environment are not considered to come about through external reinforcement. Rather, equilibration involves the mutual influence of the child's cognitive operations upon one another, with appropriate experiences producing cognitive conflict as the child recognizes contradictions and gaps in the structure of his cognitive schemas. Cognitive conflict initiates a process of inner reorganization of the relevant schemas, with the resultant development of concrete operations, including number conservation.

Piaget has specified several important characteristics of the preoperational child's egocentric thought processes which must be overcome in order for number conservation to develop. One of these, centration, involves the tendency to focus on a single striking perceptual feature of a set of objects while ignoring other equally important compensating features, which results in distorted thinking. The concrete operational child is able to decenter his thinking by taking into account all of the essential features of what he observes, so that mutually compensating changes are recognized as counteracting one another. For example, if two rows of objects are placed in positions of one-to-one correspondence, and one row is extended while the other is left intact, the preoperational child will assert that the extended row now has more elements because it is longer. Because he focuses solely on the length of the transformed row, and ignores the compensating change of decreased density, he fails to recognize the logical necessity of invariance in the number of elements. Irreversibility is another distorting quality of the preoperational child's thought processes. This refers to the child's failure to recognize that many physical transformations carried out on objects can be reversed or undone. In the extension of a row of objects, the spatial transformation can be reversed by replacing the objects in their original position of one-to-one correspondence with the untouched row. The concrete operational child's ability to decenter his thinking and his awareness of the reversibility of physical displacements enable him to free his thinking from the distorting effects of misleading perceptual cues in tasks designed to assess number conservation.

Gruen (1) conducted a number conservation training study in which he compared a procedure based upon a cognitive-conflict hypothesis with a procedure based upon a reinforcement hypothesis. The reinforcement hypothesis was translated into knowledge of results after counting. *E* confirmed an accurate count of a set of objects, corrected an inaccurate account, and rewarded overall performance with trinkets. If an *S* counted correctly, he was told: "That's right." If *S* counted incorrectly, *E* told him the correct number. The procedure for the cognitive-conflict group was similar to that for the reinforcement group, except that after a row was transformed, *E* removed an object from the center portion of the row that *S* believed contained more objects. *E* continued to remove objects from this row until *S* changed his response. *S* was given this procedure with the rows being lengthened or shortened on different trials, with a conflict being produced between the operation of transformation and subtraction. Gruen also gave half of his *Ss* in both experimental groups and in the control group verbal pretraining designed to induce them to discriminate between length and number. The results showed that the conflict-plus-verbal pretraining group made significant gains in conservation when compared to the control group with no pretraining. Similar results were not obtained for the reinforcement group. In general, the conflict *Ss* outperformed the reinforcement *Ss*.

Kaplan (2) used a number conservation training technique that combined the processes of reversibility and reinforcement. His procedure was programmed to proceed in short steps, leading each child from simple to more complex examples. The procedure consisted of placing two sets of pegs in exact correspondence on a pegboard, and then expanding the length of one set by moving its pegs in short steps toward the edge of the board, while the other set remained stationary. The number of pegs in each set progressed from one to six, and the end peg of the expanded set always remained stationary when the sets contained two or more pegs each. After each move of a peg in the expanded set (the pegs were moved two holes at a time), *S* was asked if the number of pegs in the two sets was still the same (*Q*). *S* was told whether he was correct or not after each response to *Q*. After the expanded row had been extended to its maximum length, *S* was asked to return the pegs to their original position. Throughout the training period, whenever *S* was unable to recognize that the pegs in the two sets were of equal number, he was given one of two prompts to help him become aware of the equality. These prompts took the form of asking *S* to count the pegs in the two rows, or of reversing the transformation by placing the pegs in the expanded row back in their original positions. The results of this study showed that on an immediate



posttest for conservation, *Ss* in the training group outperformed *Ss* in a control group. On a delayed posttest, three weeks later, the experimental group did slightly better than the control group, but the differences were not significant.

The purpose of the present study was to examine the effects of a number conservation training procedure that combined certain aspects of the separate procedures used by Gruen (1) and by Kaplan (2). The present training procedure involved the cognitive-conflict aspect of Gruen's procedure and the reversibility aspect of Kaplan's procedure. Another purpose of the present study was to examine the issue of whether or not the creation of cognitive conflict is a critical element in inducing *Ss* to conserve number. This was accomplished by having two training groups which both observed the addition-subtraction and reversibility operations, but with one group being confronted with a length discrepancy between the two sets of objects while the other group was not. In the conflict training procedure, one set was transformed while the other set remained stationary, which produced the standard perceptual discrepancy between the sets. In the nonconflict training procedure, both sets of objects were transformed simultaneously, which exposed *S* to the operations of addition-subtraction and reversibility without any discrepancy in spatial correspondence being produced. Comparison of the relative effects of these two training procedures was designed to provide evidence as to whether observation of these two operations alone, in the absence of a perceptual discrepancy, can be effective in inducing number conservation. Finally, the present study examined the transfer of direct training on number conservation to the developmentally more advanced task of substance conservation, which involves continuous rather than discontinuous quantity.

## B. METHOD

### 1. Subjects

*Ss* in the presented study were 18 mentally retarded students attending a prevocational training center while living at home. *Ss* were selected on the basis of a test for understanding of "more" and "less," and pretest score on a conservation of number task. Only *Ss* who demonstrated an operational understanding of more and less, but failed to demonstrate complete number conservation, were included in the study.

The total *S* sample was divided into three groups of six *Ss* each, with the groups equated as closely as possible for WAIS *IQ*, Sanford-Binet mental age score, and number conservation pretest score. The total *S* sample ranged from 43 to 69 in *IQ*, from 5-5 to 7-9 in *MA*, and from zero to 12 (18 points

possible) on the number conservation pretest. After the Ss had been assigned, the three groups were randomly assigned to one of two training conditions or to the control condition. The means and SDs of the three groups for *IQ* and *MA* are presented in Table 1.

TABLE 1  
MEANS AND SDs OF *IQ* AND *MA*  
FOR THE THREE SUBJECT GROUPS

Group	<i>IQ</i> <sup>a</sup>	<i>MA</i> <sup>b</sup>
Cognitive-conflict training		
Mean	55.50	6.54
SD	7.55	.67
Nonconflict training		
Mean	57.30	6.37
SD	5.09	.69
Control		
Mean	55.30	6.63
SD	9.94	.77

<sup>a</sup> *IQ* scores were obtained with use of the Wechsler Adult Intelligence Scale.

<sup>b</sup> *MA* scores were obtained with use of the 1960 L-M Stanford-Binet Test.

## 2. Materials

Eleven red and 11 white poker chips were used in the test for "more" and "less," and also in the number conservation test. A pegboard with two parallel rows of holes at  $\frac{1}{2}$  inch intervals, separated by two inches in the center portion of the board, and six white and six red pegs were used in the two training tasks.

## 3. Procedure

Ss were pretested for understanding of more-less, number conservation, and substance conservation. After the total group had been divided into three subgroups equated for pretest performance, five weeks of training were carried out with the two treatment groups. The control Ss were not interacted with by *E* during this period. In the week following the training period, all three groups were given posttests for number conservation and substance conservation.

*a. More-less test.* *S* was presented with two stacks of poker chips containing equal numbers of chips. After being instructed to count the chips in each stack, *S* was asked: "Does this stack (*E* points to one) have more chips, the same number of chips, or less chips than that one (*E* points to the other)?" After *S* responded, *E* subtracted one chip from the stack initially referred to and repeated the question (referred to hereafter as *Q*). Then *E* replaced the

chip and repeated Q again. *E* next added a chip to the initial stack, asked Q, removed the chip, and repeated Q again. This procedure was repeated until *S* performed without error on five consecutive items or failed on 12 items. The number of chips initially contained in each stack varied as follows on successive trials: 4, 6, 5, and 7. This sequence was repeated three times or until *S* achieved five successive errorless trials. Success on this test demonstrated operational understanding of "more" and "less" where length was not a relevant variable.

*b. Number conservation test.* This test consisted of three types of spatial transformation, with each transformation being carried out with use of three quantities of poker chips (9, 10, and 11 chips) in parallel rows. In each task, *S* was presented with two parallel rows of poker chips, positioned in one-to-one correspondence with the chips approximately two inches apart within each row. *S* was asked Q with the rows in the position of one-to-one correspondence, and after he agreed to the equality of number in the two rows, the relevant transformation was carried out. Q was then repeated, and *S* was asked to explain the basis on which he arrived at his judgment. The three transformations carried out were as follows:

*Transformation 1.* After Q was asked with the rows in the original position, the chips in one row were contracted so that its length was about one-half that of the other row, which was left untouched.

*Transformation 2.* After Q was asked with the rows in the original position, the chips in one row were expanded so that its length was approximately twice that of the other row, which remained intact.

*Transformation 3.* After Q was asked with the rows in the original position, the chips in one row were subdivided into two groups which were displaced outward so that their ends extended beyond the ends of the intact row.

On each of the nine tasks, the response to Q (judgment) and the explanation of the answer (justification) were evaluated separately. If *S* responded to Q by denying the equality of elements in the two rows after the transformation, his judgment was scored zero. Each judgment and each justification for each of the nine tasks could be scored zero, 1, or 2, which resulted in a range of possible scores on the conservation test from zero to 18.

*c. Substance conservation test.* *S* was presented with two balls of clay, equal in weight and volume, and asked Q. After *S* agreed that both balls contained the same amount of clay, each of three transformations of shape was carried out in three successive tasks. These three transformations were as follows: (a) One ball was changed into a hot dog shape while the other was left in a ball. (b) One ball was changed into a pancake while the other was

left in a ball. (c) One ball was broken into six small balls while the other was left in a large ball. After each transformation, Q was asked, and S was requested to explain the basis of his judgment. Each judgment and each justification were scored zero or 1, resulting in a possible range of scores from zero to 6.

d. *Training procedures.* Ss in the two training groups received two 30-minute training sessions per week for five consecutive weeks. In both training conditions, each session included two transformations in which the addition-subtraction operation was carried out, and two transformations in which the reversibility operation was carried out. A pegboard with two sets of six pegs each was used for the demonstrations involving both procedures.

(1). *Addition-subtraction training.* Six pegs were placed in each row, in positions of one-to-one correspondence, and S was asked to count the number of pegs in each row, to establish their equality. In the cognitive-conflict training condition E then expanded one row to twice the length of the other, which remained intact, so that one empty hole separated each peg. S was asked the following Q: "Now which row has more pegs, this row or that row, or do they have the same number?" After S responded that the expanded row had more pegs, E began adding one peg at a time to empty holes, beginning at the center of the row and working alternately toward both ends. After each addition, Q was repeated, and the additions were continued until S changed his response to Q, or all the empty holes were filled. The two rows of six pegs each were then placed in positions of one-to-one correspondence, with one empty hole separating each peg. After S confirmed the equality of number in the two rows by counting, E contracted one row by placing the six pegs in adjacent holes and asked Q. When S had stated that the contracted row had fewer pegs, E began subtracting pegs from the longer row, beginning at the center and working alternately toward both ends. After each subtraction, Q was repeated, and this procedure was continued until S changed his response to Q, or all the pegs except the two end ones had been removed. Both the addition and subtraction procedures were designed to produce cognitive conflict in S through creating a disparity between the length and the density of the longer row. By forcing S to attend to the density dimension while the length dimension remained constant, the importance of compensation of relations between the length and density dimensions was demonstrated. In the non-conflict training condition the procedures carried out were identical to those in the cognitive-conflict condition, except that the expansion-contraction and addition-subtraction operations were conducted simultaneously for both rows. Thus, the two rows always remained in positions of one-to-one correspondence



so that no perceptual disparity was ever produced between the rows. This training procedure demonstrated the addition-subtraction operations without producing cognitive conflict, since no unilateral discrepancy between the length and density dimensions occurred.

(2). *Reversibility training.* *E* initially placed two pegs in each row, in one-to-one correspondence in adjacent holes, one-third of the distance from *E*'s left edge of the board. In the cognitive-conflict situation, *E* then began extending one row by moving the center peg one hole at a time, asking *Q* after each move. This was continued until *S* denied equality of number for the two rows, or until the edge of the board to *E*'s right was reached. As soon as *S* denied equality, the peg was moved one hole back toward the starting point and *Q* repeated. The reversals were continued until *E* again responded to *Q* with an equality statement. This procedure was repeated with increasing numbers of pegs until each row contained six pegs. The end peg of the expanded row, to *E*'s left, always remained stationary, and when three or more pegs were involved, the peg nearest the stationary one was always moved next. In reversing the extension, the peg furthest from the stationary end peg was always moved to the closest empty hole from the extended end of the row. In the nonconflict training condition, the procedures were identical except that both rows were expanded and contracted simultaneously. Thus, *S* received experience with the reversibility operation without being faced with a spatial discrepancy between the two rows.

### C. RESULTS

The means and *SD*s of the three *S* groups on the number conservation pretest and posttest are presented in Table 2. Also included in Table 2 are the means and *SD* of the three *S* groups on the substance conservation pretest and posttest.

The direct training effects of the two training procedures were examined by an analysis of variance of the number conservation pretest and posttest scores summarized in Table 2. This analysis produced significant variances associated with the two tests ( $F = 8.98$ ,  $df = 1/15$ ,  $p < .01$ ), and with the interaction between the two tests and the three *S* groups ( $F = 4.31$ ,  $df = 2/15$ ,  $p < .05$ ). Examination of Table 2 shows that the significant test variance resulted from higher performance on the posttest than on the pretest by the two training groups. The critical difference procedure recommended by Lindquist (3) was used to conduct comparisons among the six separate means involved in the significant tests  $\times$  *S* groups interaction. A critical difference of 6.41 ( $p = .05$ ,  $df = 15$ ) was obtained for these comparisons, and application



TABLE 2  
MEANS AND *SDs* OF PRETEST AND POSTTEST PERFORMANCE ON THE  
NUMBER CONSERVATION AND THE SUBSTANCE CONSERVATION TESTS

Group	Number conservation <sup>a</sup>		Substance conservation <sup>b</sup>	
	Pretest	Posttest	Pretest	Posttest
Cognitive-conflict training				
Mean	2.83	12.00	.67	1.50
<i>SD</i>	4.16	6.32	1.10	1.89
Nonconflict training				
Mean	2.16	10.66	.00	1.00
<i>SD</i>	3.38	5.73	.00	1.53
Control				
Mean	4.17	2.17	1.90	1.00
<i>SD</i>	3.67	1.95	1.07	1.41

<sup>a</sup> Maximum number conservation test score = 18.

<sup>b</sup> Maximum substance conservation test score = 6.

of this figure to the means presented in Table 2 shows that posttest performance was significantly higher than pretest performance in both treatment groups, but not in the control group. Comparisons among the groups revealed that the three pretest means did not differ significantly; however, on the posttest both treatment group means were significantly higher than the control group mean, while the two treatment group means did not differ significantly.

The transfer of training effects were tested by an analysis of variance of the substance conservation pretest and posttest scores summarized in Table 2. This analysis produced no significant variances, either for the main effect of the two tests, the main effect of the three *S* groups, or for the interaction between the two tests and the three *S* groups.

#### D. DISCUSSION

The analysis of direct training effects showed that both training procedures were effective in inducing number conservation responses, at least on a short-term basis. Both training groups made significantly large increases over their pretest performance level, and both also significantly exceeded the control group on the posttest. Further examination of the separate components of the number conservation test, however, revealed that the positive training effects were not uniform among all of the experimental *Ss*. In the cognitive-conflict group one *S* continued to perform at a preoperational level on all the test items, while two others made conservation judgment responses but failed to give acceptable justifications. In the nonconflict group, three *Ss* made only five of nine conservation judgments and also failed to give acceptable justifications.

fications. Thus, while the two training procedures both produced significant changes in group means, half of the individuals in each group were still not completely trained after 10 sessions.

The present results suggest that cognitive conflict, as operationally defined by Gruen, is not essential in order to produce performance changes on the number conservation task. The five Ss in the cognitive-conflict group who showed changes in their judgment responses on the number conservation test did have a slightly higher proportion of positive responses (43 of 45) than did the Ss in the nonconflict group (41 of 54). However, this difference was not sufficiently great to exceed chance variation. On the basis of these findings it must be tentatively concluded that experience with the operations of addition-subtraction and reversibility alone, even in the absence of cognitive conflict, is sufficient to induce number conservation in some Ss.

The analysis of the transfer of training results clearly indicates that the number conservation training did not modify performance on the substance conservation task. Since substance conservation has been clearly shown to be developmentally more advanced than number conservation, this lack of transfer is not surprising. In terms of Piaget's theory this would be explained on the basis that the conservation schema had only partially assimilated discontinuous quantity, or that the amount of accommodation of the schema necessary to include continuous quantity was too great. Incomplete assimilation of discontinuous quantity by the conservation schema would necessitate a more extended training program with that type of quantity. If perfect performance for conservation of discontinuous quantity still does not enable S to conserve continuous quantity, then specific conservation training with that type of material would be indicated.

The present study was exploratory in nature and was restricted to mentally retarded Ss. Further research on the effects of such training on a broader range of S populations should be carried out to provide comparative data. Evidence regarding the relative number of training sessions required by intellectually normal *versus* retarded Ss of comparable mental ages to acquire various types of conservation would provide clarification of the relative rates of cognitive development of these S groups. It will also be necessary to examine the stability of effects of this training procedure by expanding the design to include delayed posttests over a period of several months.

#### REFERENCES

1. GRUEN, G. Experiences affecting the development of number conservation in children. *Child Devel.*, 1965, **36**, 963-975.

2. KAPLAN, J. D. Teaching number conservation to disadvantaged children. Unpublished Doctoral dissertation, Columbia University, New York, New York, 1967.
3. LINDQUIST, E. F. Design and Analysis of Experiments in Psychology and Education. Boston: Houghton Mifflin, 1953.
4. PIAGET, J. The Psychology of Intelligence. New York: Harcourt Brace, 1950.

*Human Development Center*  
*Eastern Washington State College*  
*Cheney, Washington 99004*



## SELECTION OF NaCl SOLUTIONS BY SODIUM-DEPRIVED MONGOLIAN GERBILS IN RICHTER-TYPE DRINKING TESTS\*<sup>1</sup>

*Pavlovian Research Laboratory, Veterans Administration Hospital, Perry Point, Maryland;  
and Department of Psychology, Oklahoma State University*

---

JOSEPH W. CULLEN<sup>2</sup> AND ARTHUR E. HARRIMAN

---

### SUMMARY

The study investigated whether gerbils would compensate for a dietary-induced sodium deficiency by freely ingesting NaCl solutions. Accordingly, adult Mongolian gerbils were fed either a sodium-deficient diet ( $N = 30$ ) or a sodium-replete ration ( $N = 30$ ) for 30 days. Then, each group was divided into three subgroups (five males and five females per subgroup). One sodium-deprived subgroup and one sodium-replete subgroup were assigned to Richter-type drinking tests with each of three molar NaCl concentrations (.05 M, .15 M, and .30 M) *versus* water. Although the sodium-depleted animals did not lose weight, a dramatic alopecia obtained, suggesting an effective sodium deficiency. Correspondingly, the sodium-deficient gerbils demonstrated greater acceptance of, and consumed more, NaCl than the sodium-replete animals; and both these measures were concentration dependent. These results in light of previous work by the authors suggest that the gerbil is more behaviorally adaptive under conditions of gradually increasing sodium need than under the more precipitous requirements induced by adrenalectomy or formalin injections.

---

### A. INTRODUCTION

Typically, animals depleted for sodium by dietary restriction, bilateral adrenalectomy, or formalin injections show increased appetite for sodium-bearing compounds (1, 2, 3, 7). Unlike other species tested thus far, Mongolian gerbils (*Meriones unguiculatus*) show no evidence of sodium hunger,

---

\* Received in the Editorial Office on January 22, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The authors wish to thank Dr. Victor Schwentker, Executive Director, The West Foundation, Brant Lake, New York, for providing the gerbils used in this study.

<sup>2</sup> Requests for reprints should be sent to the first author at the address shown at the end of this article.



whether the sodium is provided as a liquid or in solid form, after bilateral adrenalectomy (2). Also, gerbils rendered acutely sodium deficient via subcutaneous formalin injections do not exhibit the prompt appetite for sodium observed with laboratory rats. In contrast, their compensatory sodium selection is delayed from 48-96 hours, is not dose related, and does not occur with repeated testing (1). In the present study, gerbils were depleted for sodium through dietary restriction to determine whether the consequent hyponatremia would arouse sodium appetite despite the ineffectiveness of other procedures.

## B. MATERIALS AND METHOD

The subjects were 30 male and 30 female Mongolian gerbils (*Meriones unguiculatus*) of the Tumblebrook strain which were 172 days of age at the outset of the study. The animals were maintained in a Percival environmental chamber in which the ambient temperature was 25.0 C, and the relative humidity was 50.0%. The chamber was lighted from 8:00 a. m. to 8:00 p. m. daily. All animals were individually housed in cages with 5/8-inch mesh flooring elevated above drops to reduce coprophagia. Each cage held a metal food cup and a pair of Wahmann 140-ml calibrated drinkers. Food and drinking fluids were provided *ad libitum* throughout the study. The gerbils were fed either a sodium-deficient mash or a sodium-replete mash. The basic food mixture consisted of items, except for a substituted salt mix, that maintained Tumblebrook-strain gerbils during dietary self-selection (4). Items were mixed in amounts proportional to mean daily intakes of foodstuffs from the different categories during self-selection as follows: "Vitamin-free" soybean case in, Skidmore (25.3%); sucrose (53.4%); Hartroft-Eisenstein sodium-free salt mix, General Biochemicals (4.8%); and Wesson oil (14.5%). The sodium-deficient mash was formed by addition of 1.0% Alphacel (Nutritional Biochemicals) to the basic mixture, and 1.0% NaCl was added to the basic mixture to make the sodium-replete mash. Drinking fluids were either distilled water or NaCl solutions in which the solute was of reagent grade.

Initially, the 60 gerbils were divided randomly into two equally sized groups, with the restriction that both groups held the same number of each sex. During a 30-day pretest, one group was fed the sodium-deficient mash, and the other group was fed the sodium-replete mash. Food cups and drinkers were replaced with clean containers holding fresh materials every 24 hours. After the pretest, each diet group was divided into three subgroups for additional testing. Assignments were again made randomly, with the restriction that each subgroup held five males and five females.

The three sodium-deprived subgroups and the three sodium-replete subgroups were then tested for acceptance of NaCl solutions in 10 successive 24-hour, Richter-type drinking tests (NaCl solution opposite distilled water). Each sodium-deprived subgroup was paired with a different sodium-replete subgroup, and each pair of subgroups was tested with a different NaCl solution drawn from the following group of molar concentrations: .05 M, .15 M, and .30 M. The pair of drinkers was replaced after each Richter-type test with clean drinkers holding fresh fluids, and the fluids were switched in position from test to test. Diets for the sodium-deprived gerbils and for the sodium-replete gerbils remained unchanged from the pretest.

## C. RESULTS

### 1. Signs of Sodium Depletion

Body weights were not significantly affected by withdrawal of sodium from the diet during the 30-day pretests. Mean body weight for the sodium-deprived gerbils was 73.5 g (*SD* 8.4) at the start and was 76.3 g (*SD* 11.4) at the end of the pretest. In comparison, mean body weights for the sodium-replete gerbils at the start and end of the pretests were 77.4 g (*SD* 10.1) and 78.9 g (*SD* 10.2), respectively. A simple analysis of covariance (11) for dissimilarity in rate of weight change between the two diet groups showed no significant difference ( $F = .40$ ,  $1/57$  *df*,  $p > .20$ ).

The single externally observable effect from the feeding schedules in the pretest was an extensive alopecia in the sodium-deprived gerbils which commenced after 14 days of sodium restriction. By the end of the pretest, all but one of the gerbils exhibited the symptom, while none of the sodium-replete gerbils gave any sign of the condition. Alopecia was noted first in the abdominal region. As sodium restriction continued, the area of hair loss typically increased over the ventral surfaces and, eventually, extended from thorax to groin. Also, most of the sodium-deprived gerbils lost hair from the lower flanks and inner thighs during the final pretest days. Further, at the end of the pretest, nine males and six females in the group of 30 animals exhibited loss of hair from the crown of the head and from the shoulders and/or buccal region.

### 2. Signs of Recovery

A slight increase in mean body weight by the sodium-depleted animals (2.8 g, *SD* 4.8) during the Richter-type tests differed insignificantly from a similar increase by the sodium-replete animals (.9 g, *SD* 4.2), as tested by a

one-way analysis of variance ( $F = 1.28$ ,  $2/57$  *df*,  $p > .20$ ). There was, however, marked recovery from alopecia in 27 of the sodium-deprived gerbils. Dense, light-colored hair appeared on the ventral surfaces after the fifth Richter-type test and, thereafter, on the lateral surfaces. Traces of new hair were present on the head and shoulders in some of the animals by the end of testing, but hair was not fully restored to these areas until some days after the experiment. Slow recovery from alopecia was shown by three males in the .30 M NaCl subgroup.

### 3. Solution Acceptance

Mean percentages of acceptance for the three molar NaCl concentrations during the series of 10 Richter-type drinking tests with sodium-deprived and sodium-replete subgroups are presented in Table 1. The percentages of acceptance were treated by means of an analysis of variance for a  $p \times q$  factorial experiment (11) and indicated an insignificant interaction ( $F = .98$ ,  $2/54$  *df*,  $p > .20$ ). Also, the sodium-deprived animals showed a significantly greater acceptance of the NaCl solution than did the sodium-replete animals ( $F = 5.66$ ,  $1/54$  *df*,  $p < .05$ ). Extents of acceptance differed as a function of concentration ( $F = 10.84$ ,  $2/54$  *df*,  $p < .01$ ). Posttest comparisons by the Newman-Kuels procedure (11) indicated the following order of acceptance: .05 M  $>$  .15 M = .30 M ( $p < .05$ ,  $2/54$  *df*). In a second statistical analysis, percentages of NaCl acceptance in the first five tests and second five tests were contrasted by a two-factor experiment with repeated measures on one factor (11). No significant difference obtained between sets of tests, and other results repeated findings from the earlier analysis. Within treatment comparisons of NaCl acceptance at the three molar concentrations were performed by one-way analyses of variance. The sodium-deprived gerbils accepted the .15 M concentration in significantly greater extent than did the sodium-replete animals ( $F = 6.25$ ,  $1/18$  *df*,  $p < .025$ ); differences at the other concentrations proved insignificant ( $p > .05$ ). Acceptance of the NaCl concentrations was unrelated to sex factors ( $p > .20$ ,  $1/24$  *df*) for either the sodium-deprived gerbils ( $F = 1.43$ ) or the sodium-replete gerbils ( $F = 1.84$ ).

### 4. Actual NaCl Intakes

Table 1 also presents mean intakes of NaCl (mg/100 g body weight/test) for the sodium-deprived and sodium-replete gerbils. Statistical treatment of the NaCl intakes summarized in Table 1 was carried out with use of the analysis of variance for a  $p \times q$  factorial experiment (11). Results from

TABLE 1

ACCEPTANCE OF MOLAR NaCl SOLUTIONS AS PERCENTAGES OF TOTAL FLUID INTAKES  
AND CONSUMPTION OF NaCl (mg/100 G BODY WEIGHT) BY SUBGROUPS  
(SUBGROUP  $N = 10$ ) OF SODIUM-DEPRIVED AND SODIUM-REPLETE  
MONGOLIAN GERBILS DURING RICHTER-TYPE DRINKING TESTS

Subgroups	Molar NaCl concentration					
	.05 M		.15 M		.30 M	
	Mean	SD	Mean	SD	Mean	SD
<i>Acceptance of NaCl concentration (%)<sup>a</sup></i>						
Na-deprived	41.0	25.6	29.0	20.8	14.3	6.2
Na-replete	32.2	20.7	10.4	10.2	10.4	6.8
<i>NaCl intakes in mg/100 g body weight/test</i>						
Na-deprived	10.7	8.0	29.5	26.3	26.2	11.7
Na-replete	7.1	4.1	8.3	7.0	15.1	13.3

<sup>a</sup> % acceptance = NaCl solution/NaCl solution + distilled water  $\times 100$ .

the analysis were complementary to those obtained in treatment of the solution acceptance data. NaCl intakes were significantly higher for the sodium-deprived gerbils over the three concentrations ( $F = 11.29$ ,  $1/54$  df,  $p < .01$ ), with no significant interaction ( $F = 2.05$ ,  $2/54$  df,  $p < .10$ ). Further, NaCl intakes differed as a function of the concentration level at which the solutions were offered ( $F = 4.22$ ,  $2/54$  df,  $p < .05$ ). The Newman-Kuels procedure (11) was used to rank NaCl intakes as follows:  $.05 \text{ M} < .15 \text{ M} = .30 \text{ M}$  ( $p < .05$ ,  $2/54$  df). Also, within treatment comparisons by one-way analyses of variance indicated that the sodium-deprived group ingested significantly more NaCl in tests with the  $.15 \text{ M}$  concentration ( $F = 6.09$ ,  $1/18$  df,  $p < .025$ ). Differences between sodium-deprived and sodium-replete gerbils in tests with the other concentrations were insignificant ( $p > .05$ ).

#### D. DISCUSSION

The present findings are consistent with other reports showing that the pattern of NaCl selection in the Mongolian gerbil differs from that in laboratory rats and, possibly, in other rodents. Although rats may ameliorate a sodium deficiency induced by bilateral adrenalectomy through increased intake of NaCl (e.g., 3, 7), the gerbil does not (2). Also, although formalin injections reliably produce prompt and marked appetite for sodium in the rat, these injections evoke a transient and severely delayed sodium appetite in gerbils (1). Nevertheless, alopecia, which is a diagnostic sign of sodium deficiency in rats (6), dogs (10), and cattle (8), became severe after a relatively short period of sodium restriction. In comparison to gerbils, which lost hair after the 14th pretest day, young rats, fed a low sodium diet, showed loss of hair after eight to 10 weeks, while the sign appeared in mongrel



puppies after four weeks and in cattle after several months of sodium depletion. Despite indication of a readily developing deficiency in gerbils fed a sodium-free diet, Table 1 shows that these gerbils differed from control animals only in Richter-type tests with a .15 M NaCl concentration.

The findings summarized in Table 1 suggest that aversion to the taste of NaCl solutions, degrees of NaCl dilution, and levels of sodium need may have contributed to NaCl acceptance in the different Richter-type tests. Harriman (5) reported need free gerbils moderately rejected .05 M NaCl solutions and strongly rejected higher concentrations in single Richter-type tests. Presumably, acceptance for NaCl solutions should have been greater at all concentrations in the sodium-deprived than in need-free gerbils. Comparison of solution acceptance values with those for NaCl intakes (Table 1) reveals that sodium-deprived gerbils regulated NaCl intakes from the hypertonic solutions despite increased taste aversion with increased concentration intensity. Thus, NaCl solution acceptance at the .30 M concentration, about half that shown for the .15 M concentration led to equivalent NaCl intake (mg/100 g body weight). Although the .05 M concentration was significantly more acceptable than were the hypertonic concentrations, the solution-deprived gerbils did not sufficiently override characteristic aversion to NaCl solutions to approach the other sodium-deprived groups in NaCl intake. Despite significantly lower consumption of NaCl (mg) at the .05 M concentration, the sodium-deprived gerbils showed good recovery from alopecia during testing. Possibly, noxious stimuli associated with sodium deficiency were balanced in the different tests by need factors, on one hand, and inputs derived from aversive taste stimuli and fluid tolerance limits, on the other. In this analysis, the .15 M concentration occupied an optimal position. This concentration supplied a relatively large amount of NaCl even at a low level of solution acceptance without involving upset of the fluid balance.

Further, patterns of food and water regulation in gerbils, indigenous to xeric regions of Mongolia and China, presumably evolved to meet enduring limitations upon availability of food and water. Thus, gerbils may be less prepared physiologically for sudden and rare changes in the internal environment, as from adrenalectomy and formalin dosages, than for more gradually developing circumstances in the habitat. Hence, stressors acting acutely on the sodium metabolism may not be met effectively or promptly, while behavioral adjustments in the absence of dietary sodium are more manageable. The interpretation is supported by the sodium-reservoir hypothesis advanced by Stricker and Wolf (9) who find hyponatremia and hypovolemia as emerging coordinately in sodium deficiency. Each condition is presumed



to be separately capable of evoking sodium appetite. In line with the hypothesis, sodium restriction during the 30-day pretest elicited hypovolemia and hyponatremia which, in concert, caused the increased sodium appetite observed in the Richter-type tests. Changes in the internal environment of adrenalectomized gerbils and of gerbils subjected to formalin injections, however, contrast with those affected by insidious dietary depletion. Adrenalectomized gerbils show hyponatremia, without a concurrent and physiologically effective hypovolemia, because hypovolemia arouses sodium appetite solely through increased aldosterone secretion, and there is no extra-adrenal route (9). Also, neither the hypovolemia nor the hyponatremia present in gerbils shortly after injection of formalin evoked sodium appetite until the stressor overcame, after several days, a stubborn resistance to electrolytic imbalance in this desert species (1).

### REFERENCES

1. CULLEN, J. W. Sodium intake in the Mongolian gerbil (*Meriones unguiculatus*) consequent to subcutaneous formalin injections. *Psychon. Sci.*, 1972, **26**, 279-282.
2. CULLEN, J. W., & SCARBOROUGH, D. E. Behavioral and hormonal prophylaxis in the adrenalectomized gerbil (*Meriones unguiculatus*). *Hormones & Behav.*, 1970, **1**, 203-210.
3. FREGLY, M. J. Specificity of the sodium chloride appetite of adrenalectomized rats: Substitution of lithium chloride for sodium chloride. *Amer. J. Physiol.*, 1958, **195**, 645-653.
4. HARRIMAN, A. E. A comparative study of food and water regulation by laboratory rats and Mongolian gerbils maintained on identical self-selection of diet schedules. *Amer. Midl. Natur.*, 1969, **82**, 157-162.
5. ———. Preferences by Mongolian gerbils for solutions of sugars, acids, and salts in Richter-type tests. *J. Gen. Psychol.*, 1970, **83**, 79-86.
6. ORENT-KELLES, E., ROBINSON, A., & MCCOLLUM, E. V. The effects of sodium deprivation on the animal organism. *Amer. J. Physiol.*, 1937, **119**, 651-661.
7. RICHTER, C. P. Total self-regulatory functions in animals and human beings. *Harvey Lectures*, 1942, **38**, 63-103.
8. SMITH, R. M. *The Physiology of the Domestic Animals*. Philadelphia: Davis, 1890. P. 192.
9. STRICKER, E. M., & WOLF, G. Behavioral control of intravascular fluid volume: Thirst and sodium appetite. *Ann. N. Y. Acad. Sci.*, 1969, **157**, 553-568.
10. TURPEINEN, O. Studies on sodium deficiency. The effects of sodium deprivation on young puppies. *Amer. J. Hyg.*, 1938, **28**, 104-110.
11. WINER, B. J. *Statistical Principles in Experimental Design* (2nd ed.). New York: McGraw-Hill, 1971. Pp. 431-445, 518-539, 775-780.

Grants Associate Program  
Division of Research Grants  
National Institutes of Health  
Bethesda, Maryland 20014



## A COMPARISON OF RESPONSE RATES IN RESPONSE-TERMINATED AND TIME-TERMINATED EXPERIMENTAL SESSIONS\*

*University of Florida and Georgia State University*

ALAN C. REPP,<sup>1</sup> SAMUEL M. DEITZ,<sup>2</sup> AND WILLIAM D. WOLKING

### SUMMARY

Three subjects were exposed to fixed-ratio schedules of reinforcement in a two-phase experiment. In the first phase, sessions were terminated after a fixed number of responses had been emitted. In the second phase, sessions were terminated after a fixed length of time (equivalent to the mean of session lengths when steady state responding occurred in Phase 1). A comparison of response rates showed higher rates for all subjects in Phase 2.

### A. INTRODUCTION

Imposed variability caused by methodological arrangements often produces misinterpretation of data (4). One procedural decision which might influence responding is made before data collection begins, but could, in itself, cause higher or lower response rates. That decision is whether to extend an experimental session until a predetermined amount of reinforcers has been earned or until a certain amount of time has passed. When fixed-ratio (FR) schedules of reinforcement are used, this decision becomes one of whether to terminate sessions after a number of responses has been emitted or after an amount of time has elapsed. Basic research using animal and human subjects has been completed by use of both procedures. The response rates obtained, however, may have been at least partially dependent on the procedure chosen, as well as on the independent variables under investigation. Applied research also utilizes both procedures, although most studies terminated sessions by time and allowed a flexible number of reinforcers to be earned.

One type of applied research which recommends both procedures is O. R. Lindsley's Precision Teaching System (1, 2, 3). In this system, sessions are

\* Received in the Editorial Office on January 22, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Now at the Georgia Retardation Center, Atlanta. Reprints may be obtained from A. C. Repp at the address shown at the end of this article.

<sup>2</sup> The second author is at Georgia State University.

concluded either with respect to time, the number of responses emitted, or a combination of both. If one method produces a different response rate from the other, researchers could plan to terminate sessions by the method most suitable for their purposes.

The purpose of the present study was to determine whether there would be different (higher or lower) response rates produced by experimental sessions ended by time or ended by a predetermined number of responses.

## B. METHOD

### 1. Subjects

Three female college sophomores, ages 19, 19, and 20, served as subjects. All attended the University of Florida, and none had ever participated in an experiment of this type before.

### 2. Apparatus

Subjects were seated individually in the experimental chamber, a room 3 feet by 5 feet (.91 m  $\times$  1.52 m). They faced a 12 inch by 12 inch (.31 m  $\times$  .31 m) stimulus panel mounted on the front wall and centered 3 feet (.91 m) above the floor. One microswitch button was mounted on the panel. Its operation required a force of approximately 20 g through a distance of 1 cm. Two rows of lights, one row of two red 7-watt lamps and one row containing one green and one yellow 7-watt lamps, along with a four-digit counter, were located above the manipulandum.

The two red lamps provided the only continuous illumination in the chamber. The green and yellow lights were correlated with stimulus conditions. A loudspeaker above the stimulus panel delivered white noise to mask extraneous sound.

### 3. Procedure

The experiment was conducted in two phases. In Phase 1, while the green light was on, each response resulted in a point added to the counter, and each 50 points resulted in one cent earned by the subject. This combined schedule can be designated FR 50 (FR 1). The session ended when 5000 responses had accumulated on the counter.

Each session began with the counter at zero. The instructions, given at the beginning of session one, were as follows: "You can change the number on that counter by pressing the button. The overhead light will be off, but the two red lights will be on so that you can see the number on the counter. Sometimes the green and sometimes the yellow light will be on; when neither

is on, the session is over. At this time you will have earned one penny for every 50 points on the counter." When a session was over, students signed a ledger on which the earned points were totaled. They were paid at the end of the experiment.

Phase 1 was in effect for a minimum of 10 sessions and was ended when steady state responding occurred. The criterion chosen required that the rates in the last 10 sessions did not differ from their mean rate by more than five percent.

During Phase 2, the yellow light was on, the green light was off, and subjects responded under the FR 50 (FR 1) schedule. In this phase, however, sessions were ended after a prespecified amount of time had passed. The length of time was figured separately for each subject and was the mean length of time spent during the 10 steady state sessions of Phase 1. When 10 consecutive sessions meeting the steady state criterion were completed, the experiment was concluded.

### C. RESULTS AND DISCUSSION

Figure 1 shows the rate of responding for all subjects in the study. The graphs show the information from all parts of the study, but only the data from the steady states of each phase will be discussed.

In the 10 sessions of steady state responding during Phase 1, subject number 1 (S-1) produced a mean response rate of 4.28 responses per second (within the criterion range of 4.07 to 4.49 responses per second). During Phase 2, the mean response rate increased to 4.63 responses per second (within the steady state range of 4.40 to 4.87 responses per second). When matched pairs of response rates from Phase 1 to Phase 2 were compared, every rate from Phase 2 exceeded its matched rate from Phase 1, and no session from Phase 2 fell below the steady state mean of Phase 1.

The steady state responding of S-2 during the first phase produced a mean of 4.18 responses per second (within the steady state criterion range of 3.97 to 4.39 responses per second). Responding increased during the second phase to a mean rate of 4.44 responses per second (within the criterion range of 4.22 to 4.66 responses per second). A comparison of matched pairs of response rates from Phases 1 and 2 found that no session of Phase 1 exceeded its matched session in Phase 2. Again, no session of Phase 2 fell below the steady state mean rate of Phase 1.

During the first phase, S-3 had a steady state mean rate of 4.41 responses per second (within the steady state range of 4.18 to 4.63 responses per second). The mean rate rose in Phase 2 to 4.72 responses per second (within



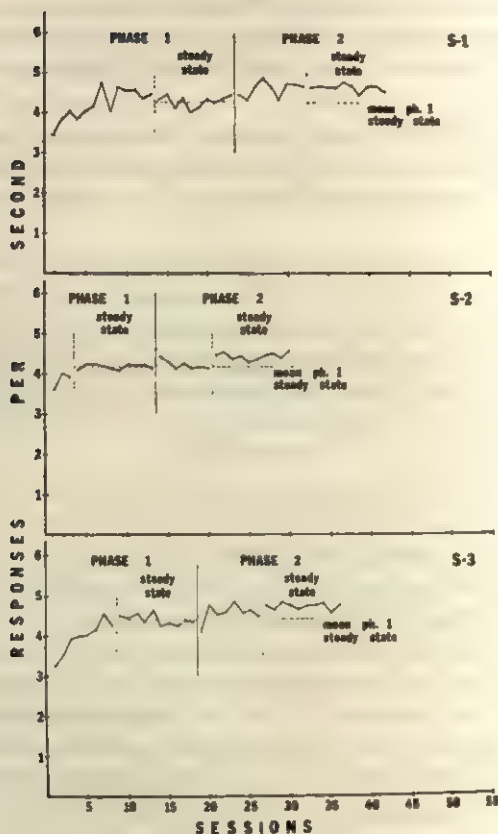


FIGURE 1

RESPONSE RATES FOR THE THREE Ss DURING PHASES 1 AND 2

Vertical dotted lines indicate the steady state sections of each phase. Horizontal dotted lines indicate the mean steady state responding in Phase 1. Phase 1 sessions were terminated after 5000 responses; Phase 2 sessions were terminated after a fixed length of time equal to the mean of session lengths during Phase 1 steady state.

the criterion range of 4.48 to 4.95 responses per second). No session of Phase 1 exceeded its matched session of Phase 2, and all response rates in Phase 2 exceeded the steady state mean response rate of Phase 1.

No differences between the two mean steady state rates were of a very large magnitude for any of the subjects. This could be due, in part, to the very high rate of responding (about four responses per second) which occurred in

both phases for all subjects. Had initial responding been somewhat lower in Phase 1, greater differences between phases may have been found.

The consistency of response rates between subjects was remarkable. Across all subjects, the mean response rate in the first phase was 4.29 responses per second, and the mean for each subject in that phase was within a five percent range of the mean (4.08 to 4.50 responses per second). The average of the means for the three subjects during Phase 2 rose to 4.59 responses per second, and all means were within the five percent range of that average (4.36 to 4.82 responses per second). In no case did any of the means from Phase 1 exceed any mean for Phase 2. Also, all three Phase 2 means exceeded the average of the means of the first phase. This consistency, as well as the difference in response rates, supports the superiority of one procedure over the other in producing higher response rates.

#### D. CONCLUSION

In some cases, procedures, rather than independent variables under investigation, can cause differences in data. In the study reported here, response rates of three Ss were compared under two different procedures. When sessions were limited to the amount of reinforcers (i.e., number of responses) possible, response rates were slightly, but consistently, lower than when sessions were limited by an amount of time but not by the number of reinforcers. If maximum response rates are preferable for a research problem, there results show the latter procedure to be preferable. Restricting time, but allowing a flexible number of reinforcers to be earned, produced higher response rates than did restricting reinforcers but allowing a flexible amount of time to elapse before the session was terminated.

#### REFERENCES

1. COUNCIL FOR EXCEPTIONAL CHILDREN. *Teach. Except. Child.* (entire issue), 1971, 3(3).
2. LINDSLEY, O. R., Ed. *Precise Behavior Facts*. Kansas City, Kansas: Precision Media, 1971.
3. PROCEDURE REPORT (PR-12). Undated computer form used for the collection of behavioral data, Behavior Research Company, Kansas City, Kansas.
4. SIDMAN, M. *Tactics of Scientific Research*. New York: Basic Books. 1960. Pp. 141-212.

*Department of Special Education*  
*Georgia Retardation Center*  
*4770 North Peachtree Road*  
*Atlanta, Georgia 30341*



## SELECTIVE EXPOSURE: AN ADDENDUM\*

*Montgomery College and The American University*

---

LEONARD L. ROSENBAUM AND ELLIOTT MCGINNIES

---

### SUMMARY

As students entered a room to hear a pro-Israel or a pro-Arab lecture, they were asked to indicate whether they sympathized with the Israelis, the Arabs, or neither side in the Middle East conflict. A greater number of students inclined to support the Israeli position attended the lecture by the pro-Israel speaker, whereas more pro-Arab students attended the lecture by the pro-Arab speaker.

---

### A. INTRODUCTION

After reviewing the status of the selective exposure hypothesis in persuasion, Freedman, Carlsmith, and Sears (2) reach the conclusion that "... whether or not selective exposure occurs depends, to a great extent, on the situation" (p. 288). In general, however, it appears that people in the natural environment tend to prefer supportive to nonsupportive information, whereas those studied under laboratory conditions frequently do just the opposite. In explaining the selectivity usually encountered outside the laboratory, these authors point to the disproportionate availability of supportive information. This paper presents the results of a field study in which two sides of an issue were equally available to a group of students. Further, the initial attitudes of the students could be related to their attendance at two public lectures where the issue was discussed.

In the spring of 1972, the Political Science Department on the Rockville Campus of Montgomery College sponsored lectures by two partisan speakers on the Middle East conflict. Dr. Elias Shoufani, a university professor, presented a pro-Arab viewpoint on April 14, and Mr. Nissim Eliad, a member of the Israeli Knesset (parliament) presented a pro-Israel viewpoint on April 26. The two lectures were publicized in the school newspaper, by posters at various locations on campus and by the instructors of various

---

\* Received in the Editorial Office on January 22, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

courses. We took the opportunity of further testing the selective exposure hypothesis in this natural setting.

## B. METHOD

### 1. Procedure

Before the beginning of each lecture *Es* waited at the entrance of the room where the talk was scheduled. Each person entering the classroom was given a short questionnaire to complete anonymously. Subjects were asked to indicate whether they were inclined to support the Israelis, the Arabs, or neither. Further, subjects were asked to indicate whether they came to the lecture voluntarily or as a course requirement. Subjects were asked to complete the questionnaire before the lecture began.

### 2. Subjects

Twenty-nine members of the audience for Dr. Shoufani's lecture and 42 of those for Mr. Eliad's lecture returned questionnaires. A combined total of fewer than six students who attended these lectures failed to complete the questionnaire. Of the total of 71 students sampled, 52 were male. There were not enough women in attendance to evaluate sex differences in selective exposure. This confirms our previous finding of a greater interest among male students in international political issues (3). Twenty-two subjects who indicated that they attended the lecture because it was a class requirement were not included in the analysis, since their responses would have no bearing on the selective exposure question.

## C. RESULTS

The numbers of pro-Arab, pro-Israeli, and neutral subjects who attended each of the two lectures are presented in Table 1. A preponderance of students who were inclined to support the Israeli position attended the lecture by the pro-Israel speaker; more pro-Arab students, on the other hand, attended the lecture by the pro-Arab speaker. More uncommitted

TABLE 1  
NUMBER OF PRO-ARAB AND PRO-ISRAELI STUDENTS AT EACH LECTURE

Audience	Speaker		Total
	Arab	Israeli	
Supported Arabs	9	4	13
Supported Israelis	5	14	19
Supported neither	11	6	17
Total	25	24	49



students also were in attendance at Dr. Shoufani's talk. These differences, when evaluated by chi square, were statistically significant ( $\chi^2 = 7.63$ ;  $p < .05$ ).

#### D. DISCUSSION

These data provide additional empirical support for the rather reasonable assumption that people in a natural setting are more inclined to expose themselves to speakers with whom they initially agree. Freedman and Sears (1) suggest that such selective exposure occurs because people tend to live in environments that disproportionately contain information supporting their attitudes. In the present situation, this assumption is not relevant, because virtually all publicity before the first lecture, in the series of two, referred to both presentations. In addition, these lectures were scheduled so that the same students could attend both. The audience members in this instance evidently sought out information with which they agreed and avoided potentially discrepant information. We have no supportable hypothesis as to why the pro-Arab speaker drew more uncommitted individuals than the pro-Israeli speaker. Perhaps these individuals felt that they lacked information regarding the Arab position. Or, perhaps this result occurred because the Arab speaker appeared on campus first.

#### REFERENCES

1. FREEDMAN, J., & SEARS, D. Selective exposure. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 2). New York: Academic Press, 1965. Pp. 58-97.
2. FREEDMAN, J. L., CARLSMITH, J. M., & SEARS, D. O. *Social Psychology*. Englewood Cliffs, N. J.: Prentice-Hall, 1970.
3. MCGINNIES, E., & ROSENBAUM, L. L. A test of the selective-exposure hypothesis in persuasion. *J. of Psychol.*, 1965, **61**, 237-240.

*Department of Psychology*  
*Montgomery College*  
*Rockville, Maryland 20850*



## LANGUAGE MEANING (GENDER SHAPING) AMONG BLIND AND SIGHTED STUDENTS\*

*Virginia Commonwealth University*

---

JOHN G. CULL AND RICHARD E. HARDY

---

### SUMMARY

Blind students showed significant differences from their sighted counterparts in assigning gender to 17 commonly used words out of 50. Language seems to help condition perception, and persons who are deprived of vision seem to have different interpretations concerning the meaning of words in the language.

---

### A. INTRODUCTION

Language plays an important part in our perception of the feeling of persons toward us and, in general, of the total environment around us (1, 2, 7). Problems relating to communication skills and sight loss have been described in the literature (3, 4, 6).

The purpose of this research was to evaluate the meaning of words within the language system for both blind and sighted students to determine whether substantially different meanings exist for the two groups.

### B. METHOD

Twenty-two totally blind students at the Virginia School of the Blind in Staunton, Virginia, and 64 sighted subjects from the Wilson Memorial High School in Fishersville, Virginia, were randomly selected from classes of senior English. These two schools were selected because of the similarity. They are both located in a rural section of Virginia. The students in both schools have a high degree of similarity in educational, cultural, and social levels, and in level of general academic achievement.

The Gender Association Survey (5) which is based on semantic differential tests was administered in May of 1971. Subjects reacted to abstract and concrete English nouns on a five-point masculinity-femininity scale where number 1 was masculine, and number 5 was feminine.

No reason for the study was given to the subjects in order that precon-

---

\* Received in the Editorial Office on January 26, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

ceived ideas about the survey could vary uncontrolled. Biographical data included age, sex, and education. None of the students used a second language at home.

### C. RESULTS

The mean scores on the Gender Association Survey for blind and sighted males and females were compared.<sup>1</sup> A standard error of the mean differences for males and females showed that both groups scored as could be expected from chance factors alone.

For purposes of cross-validation, both the blind and sighted groups were then divided into random subgroups, and the same statistical procedure was applied. Results indicated that when comparisons were made between the two groups of blind subjects, only two words were interpreted significantly different at the .05 level. The comparison between the two groups of sighted subjects in interpreted response indicated that three words were interpreted significantly different at the .05 level.

The blind and sighted groups were compared on their responses to the masculinity and femininity connotations of the English nouns on the survey. Differences at the .05 level of significance were found for 17 words, as follows: love, progress, star, knot, chair, sun, thief, work, man, defeat, cloud, smoke, root, fear, belief, battle, life.

### REFERENCES

1. JACOBOVITS, L. A., & MIRON, M. S., Eds. *Readings in the Psychology of Languages*. Englewood Cliffs, N. J.: Prentice-Hall, 1967.
2. KORZYBSKI, A. *Science and Sanity*. New York: Country Life Press, 1933.
3. MINOR, L. E. A study of the incidence of speech deviations among visually handicapped children. *New Outlook for Blind*, 1963 (Jan.), 57(1), 10-14.
4. MOOR, P. M. Blind children with development problems. *Children*, 1961 (Jan-Feb.), 1, 9-13.
5. OSOON, C. E. The Gender Association Survey. Unpublished manuscript, Institution of Communicational Research, University of Illinois, Urbana, 1967.
6. ROGOW, S. The nonverbal blind child: Two paths to speech. *New Outlook for Blind*, 1969 (Jan.), 63(1), 1-7.
7. SAPORTA, S., Ed. *Psycholinguistics*. New York: Holt, Rinehart & Winston, 1961.

Virginia Commonwealth University  
Regional Counselor Training Program  
Box 499 WWRC  
Fishersville, Virginia 22939

<sup>1</sup> Tabular material available from the authors at the address shown at the end of this article.

## EFFECTS OF GENETIC COUNSELING ON PARENTAL SELF-CONCEPTS\*<sup>1</sup>

*Departments of Pediatrics and Medical Genetics, Indiana University Medical Center;  
and Department of Neurology, Medical College of Georgia*

---

MARY ANN ANTLEY, RAY M. ANTLEY, AND LAWRENCE C. HARTLAGE<sup>2</sup>

---

### SUMMARY

Eighteen pairs of parents of children with Down's syndrome were tested on self-concept items prior to and immediately following genetic counseling. There was significant improvement on five of 20 self-concept items, with mothers improving on four items, and fathers on two items. Only one item demonstrated improvement by both mothers and fathers. Genetic counseling was found to be of value in improving self-concepts of parents of children afflicted with at least one type of heritable disease.

---

### A. INTRODUCTION

Within the last decade, there has been considerable progress in medical genetics in the determination of both causative elements and recurrence risks in a number of heritable diseases. Although there is evidence that indicates that parents of children afflicted with a heritable disease suffer serious problems in loss of self-concept and feeling of self-worth (2), there has been no corresponding study of the effects on self-concept of the delivery of genetic counseling to affected individuals or their relative family members. This study was aimed at investigating the effects of genetic counseling on a number of self-concept variables in parents of children afflicted with Down's syndrome. Down's syndrome was chosen as a model disease both because of its frequency in the population and because of its relatively predictable outcome and course (3).

---

\* Received in the Editorial Office on January 29, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This work was supported by research grant (500-838-64) from the U.S. Public Health Service and by a grant for Maternal and Child Health Project 924 from the U.S. Department of Health, Education and Welfare. The preparation of the manuscript was aided by support from the Riley Memorial Foundation.

<sup>2</sup> Direct all requests for reprints to the third author at the address shown at the end of this article.



## B. METHOD

### 1. *Subjects*

Eighteen consecutive pairs of parents who presented themselves at a large medical center for genetic counseling were studied. The parents ranged in age from 19 to 40, and were for the most part of middle socioeconomic level. Counseling was done within approximately three months of the birth of the affected child, and this counseling represented the first specific diagnosis of Down's syndrome that was given to the parents.

### 2. *Measurement Instrument*

A semantic differential self-concept scale containing 20 pairs of words was used. The scale had been previously tested and found to have split-half reliabilities in the .80's (1). The item pairs were arranged on a seven point continuum, and respondents indicated their self-perceptions by checking the point along the continuum that best described them. A score of 1 indicated a low, and a score of 7 indicated a high self-concept.

### 3. *Procedure*

The parents were asked to complete the self-concept form prior to receiving genetic counseling. They were then given genetic counseling of approximately one to two hours duration by a trained genetics counselor, and, following the genetic counseling session, they were given another form of the same instrument. Scores between pre- and postcounseling sessions were compared by a paired *t* test for each individual item, for a total of 20 comparisons.

## C. RESULTS

There was a significant improvement after counseling on a number of specific self-concept items both for men and women, with the greatest improvement manifested by women. Table 1 shows the pre- and postcounseling session levels on the 20 self-concept variables for men and women, along with significance levels associated with the magnitude of their changes. Women showed significant improvement on the four following variables: indecisive-determined, cowardly-brave, accepting-rejecting, innocent-guilty; whereas men showed significant changes on only two variables: indecisive-determined and impatient-tolerant.

## D. DISCUSSION

It appears that genetic counseling exerted greater impact on the alleviation of feelings of cowardliness, rejection, and self-blame in women than in men.

TABLE 1  
PRE- AND POSTCOUNSELING SELF-CONCEPT MEAN SCORES

Item	Precounseling		Postcounseling		$t^a$	
	Men	Women	Men	Women	Men	Women
1. Helpful-critical	5.81	5.89	6.13	6.11	.864	.776
2. Impatient-tolerant	4.73	4.78	6.00	5.17	2.433**	.789
3. Ambitious-lazy	5.88	5.88	6.06	5.61	.677	1.001
4. Immature-mature	6.00	5.61	6.25	5.83	1.291	1.290
5. Happy-unhappy	5.00	5.41	5.95	5.64	.845	.696
6. Dishonest-honest	6.06	6.50	6.38	6.67	1.779	1.001
7. Dependable-undependable	6.25	6.22	6.19	6.17	.296	.251
8. Indecisive-determined	5.63	5.33	6.38	6.00	2.534*	2.288*
9. Strong-weak	5.63	4.94	6.00	5.56	1.197	1.479
10. Cowardly-brave	6.06	4.72	6.06	5.44	.018	3.198**
11. Valuable-worthless	5.56	5.44	5.75	5.50	.510	.174
12. Foolish-wise	5.77	4.89	5.87	5.28	1.06	1.161
13. Sincere-insincere	5.93	6.39	6.13	6.11	.641	1.158
14. Kind-cruel	6.31	6.22	6.13	6.28	.613	.324
15. Accepting-rejecting	5.81	5.28	5.94	6.17	.621	3.498**
16. Inconsiderate-thoughtful	5.94	6.00	6.13	5.83	.545	.546
17. Responsible-irresponsible	6.07	6.38	6.33	6.00	1.000	1.327
18. Uncooperative-cooperative	6.31	6.22	6.25	6.11	.222	.461
19. Innocent-guilty	5.25	4.17	5.56	5.44	.892	2.389*
20. Good-bad	6.25	5.78	6.13	5.89	.415	.437

<sup>a</sup>  $t$  values reflect comparisons between pre- and postcounseling sessions for men and women separately.

\*  $p < .05$ .

\*\*  $p < .01$ .

The one change that was common to both men and women involved the indecisive-determined variable, which may well be a reflection of parents' need for structure and information about possible choices regarding their handicapped child. One possible explanation for the comparatively greater improvement of self-concept of women than men following genetic counseling may involve the comparatively greater feelings of closeness to the child by the mother and the possible guilt experienced by her, since in her societal role she is the family member most involved with caring for the handicapped child.

In order to investigate the support for this possible explanation, the pretest levels on self-concept items were compared between men and women. Women with afflicted children were found to be significantly lower than their husbands prior to counseling on such variables as strong-weak, cowardly-brave, foolish-wise, and innocent-guilty, and were also lower on these variables than were a number of control women whose children were not afflicted, tending to support this possible explanation. In any case, it appears that genetic counseling can improve self-concept of parents, especially the mothers of children affected with Down's syndrome. It may be of value to consider these findings in planning for subsequent genetic counseling activities to ensure that the genetic

counseling is aimed not only at imparting factual information relative to recurrence risks and prognostic factors for the specific disease under consideration, but also for utilizing its potential for alleviating parents' poor self-concept.

#### REFERENCES

1. HARTLAGE, L. C., & HALE, P. Self-concept decline from psychiatric hospitalization. *J. Indiv. Psychol.*, 1968, **24**, 174-176.
2. LEONARD, C. O., CHASE, G. A., & CHILDS, B., Genetic Counseling: A consumer's view. *New Eng. J. Med.*, 1972, **287**, 433-439.
3. STEIN, Z., & SUSSER, M., The preventability of Down's syndrome. *Health Serv. & Ment. Health Admin. Health Rep.*, 1971, **86**, 650-658.

*Department of Neurology*  
*Medical College of Georgia*  
*Augusta, Georgia 30902*

(Founded in 1935 by Carl Murchison)

# The Journal of PSYCHOLOGY

The General Field of Psychology

POWELL MURCHISON  
*Managing Editor*

HELEN M. DAVIS  
*Associate Editor*

## *Editorial Board*

JOHN A. ANTOINETTI  
Life Insurance Agency  
Management Association

S. HOWARD BARTLEY  
Michigan State University

ARTHUR G. BILLS  
University of Cincinnati

E. A. BOTT  
University of Toronto

LEONARD CARMICHAEL  
National Geographic Society

IRVIN L. CHILD  
Yale University

HERBERT S. CONRAD  
U. S. Office of Education

JOHN F. DASHIELL  
Florida State University

LEONARD W. DOOB  
Yale University

If this space should be unstamped, this is the regular library edition.  
But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not  
be resold for a period of five years from date of publication.

FRANK A. GELDARD  
Princeton University

ERNEST R. HILGARD  
Stanford University

JOHN E. HORROCKS  
Ohio State University

DAVID HOTHERSALL  
Ohio State University

HEINRICH KLÜVER  
University of Chicago

NORMAN L. MUNN  
University of Adelaide,  
Australia

WARREN H. TEICHNER  
New Mexico State  
University

LEE EDWARD TRAVIS  
Fuller Theological  
Seminary

JOSEPH ZUBIN  
Psychiatric Institute,  
Columbia University

## VOLUME 83

1973

Copyright © 1973 by the Journal Press  
Provincetown, Massachusetts, U. S. A.

All rights reserved

Published at 2 Commercial Street, Provincetown, Mass.

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

# AUTHOR INDEX

Adam, Everett E., Jr. ....	31	Kastner, Sheldon .....	257
Antley, Mary Ann .....	335	Klayman, Bruce E. ....	173, 201
Antley, Ray M. ....	335	Lancaster, Roy P. ....	303
Barabasz, Arreed F. ....	295	Levine, Jacob .....	67
Baumeister, Alred A. ....	57	Lewis, Hadley M. ....	163
Bean, Linda L. ....	227	Liebert, Ann M. ....	57
Bledsoe, Joseph C. ....	131	Lindgren, Henry Clay .....	117
Blick, Kenneth A. ....	277	Lorr, Maurice .....	287
Buchanan, Lyle J., Jr. ....	117	McCroskery, James H. ....	227
Cull, John G. ....	333	McGinnies, Elliott .....	329
Cullen, Joseph W. ....	315	McGlothlin, William H. ....	123
Cummings, L. L. ....	215	McManis, Donald L. ....	303
Deitz, Samuel M. ....	323	Meikle, Stewart .....	17
Ecker, Jonathan .....	67	Merlis, Sidney .....	243
Finney, Phillip D. ....	89	Paul, Robert J. ....	31
Firth, John .....	151	Pedersen, Darhl M. ....	3, 79
Foote, Russell .....	263	Permut, Steven E. ....	41
Fowler, Marguerite Gilbert .....	237	Phillips, Victor K. ....	25
Fowler, Robert L. ....	237	Rambo, William W. ....	89
Fracchia, John .....	243	Repp, Alan C. ....	323
Fracher, Jeffrey C. ....	277	Rosenbaum, Leonard L. ....	329
Gaito, John .....	151	Schmidt, Stuart M. ....	215
Geen, Russell G. ....	95	Sheppard, Charles .....	243
Gerritse, Richard .....	17	Sieeking, Nicholas A. ....	49
Greenberg, Jerrold S. ....	137	Simon, William E. ....	145
Hardy, Richard E. ....	333	Steininger, Marion .....	11
Harlow, Steven D. ....	111	Stonner, David .....	95
Harnett, Donald L. ....	215	Suziedelis, Antanas .....	287
Harriman, Arthur E. ....	315	Teubner, Johanna .....	111
Hartlage, Lawrence C. ....	335	Van De Riet, Hani .....	237
Hines, George H. ....	247	Wiggins, R. Gene .....	131
Hunter, Edna J. ....	163	Williams, D. G. ....	103
Jackson, Dorothy W. ....	251	Williams, John D. ....	111
Jamison, Kay .....	123	Wolking, William D. ....	323
Jones, Warren H. ....	89	Woodward, Jack .....	263
		Zigler, Edward .....	67



# TABLE OF CONTENTS

Developmental trends in personal space . . . . .	3
BY DARHL M. PEDERSEN	
A comparison of two kinds of dogmatism scores: Rokeach categories <i>versus</i> open-ended responses . . . . .	11
BY MARION STEININGER	
A comparison of husband-wife responses to pregnancy . . . . .	17
BY STEWART MEIKLE AND RICHARD GERRITSE	
Creativity: Performance, profiles, and perceptions . . . . .	25
BY VICTOR K. PHILLIPS	
Awareness in verbal operant conditioning: Examination of performance quality changes . . . . .	31
BY EVERETT E. ADAM, JR., AND ROBERT J. PAUL	
Cue utilization patterns in student-faculty evaluation . . . . .	41
BY STEVEN E. PERMUT	
A child-adult research form of the Pittsburgh Scales of Social Extraversion-Introversion and Emotionality . . . . .	49
BY NICHOLAS A. SIEVEKING	
Behavioral variability among retardates, children, and college students . . . . .	57
BY ANN M. LIEBERT AND ALFRED A. BAUMEISTER	
Impaired sex-role identification in schizophrenia expressed in the comprehension of humor stimuli . . . . .	67
BY JONATHAN ECKER, JACOB LEVINE, AND EDWARD ZIGLER	
Relations among sensation seeking and simulated and behavioral personal space . . . . .	79
BY DARHL M. PEDERSEN	
Some correlates of the level of constraint in a system of social attitudes . . . . .	89
BY WILLIAM W. RAMBO, WARREN H. JONES, AND PHILLIP D. FINNEY	
Reactions to aggression-related stimuli following reinforcement of aggression . . . . .	95
BY RUSSELL G. GEEN AND DAVID STONNER	
So-called "nervous habits" . . . . .	103
BY D. G. WILLIAMS	
Creativity in rural, urban, and Indian children . . . . .	111
BY JOHN D. WILLIAMS, JOHANNA TEUBNER, AND STEVEN D. HARLOW	
Brainstorming in large groups as a facilitator of children's creative responses . . . . .	117
BY LYLE J. BUCHANAN, JR., AND HENRY CLAY LINDGREN	
Drug usage, personality, attitudinal, and behavioral correlates of driving behavior . . . . .	123
BY KAY JAMISON AND WILLIAM H. MCGLOTHLIN	
Congruence of adolescents' self-concepts and parents' perceptions of adolescents' self-concepts . . . . .	131
BY JOSEPH C. BLEDSOE AND R. GENE WIGGINS	
A study of the self-esteem and alienation of male homosexuals . . . . .	137
BY JERROLD S. GREENBERG	

Age, sex, and title of therapist as determinants of patients' preferences . . . . .	145
BY WILLIAM E. SIMON	
Procedures for estimating magnitude of effects . . . . .	151
BY JOHN GAITO AND JOHN FIRTH	
The dyslexic child--Two years later . . . . .	163
BY EDNA J. HUNTER AND HADLEY M. LEWIS	
Heteromodal cueing and auditory-visual interaction: A literature review	173
BY BRUCE E. KLAYMAN	
Detection and the timing of a heteromodal cue: Auditory-visual interaction	201
BY BRUCE E. KLAYMAN	
Factor similarity of personality across private and military samples: An analysis of the Personality/Attitude Schedule . . . . .	215
BY L. L. CUMMINGS, DONALD L. HARNETT, AND STUART M. SCHMIDT	
Implications for associative processes of switching the middle of the list during serial rote learning . . . . .	227
BY LINDA L. BEAN AND JAMES H. MCCROSKERY	
Feminism and political radicalism . . . . .	237
BY MARGUERITE GILBERT FOWLER, ROBERT L. FOWLER, AND HANI VAN DE RIET	
Personal adjustment of hospital staff and their attitudes about mental illness	243
BY JOHN FRACCHIA, CHARLES SHEPPARD, AND SIDNEY MERLIS	
Motivational correlates of Pacific Islanders in urban environments . . . . .	247
BY GEORGE H. HINES	
Alienation and identity-role diffusion in late adolescence . . . . .	251
BY DOROTHY W. JACKSON	
Tolerance for the unstable and defensive role adjustment in response to sudden physical disability . . . . .	257
BY SHELDON KASTNER	
A preliminary investigation of obscene language . . . . .	263
BY RUSSELL FOOTE AND JACK WOODWARD	
Speed of motor conflict resolution as related to type of conflict and manifest anxiety . . . . .	277
BY JEFFREY C. FRACHER AND KENNETH A. BLICK	
Conservative attitudes and authoritarian values . . . . .	287
BY ANTANAS SUZIEDELIS AND MAURICE LORR	
Group desensitization of test anxiety in elementary school . . . . .	295
BY ARREED F. BARABASZ	
Training of number conservation in retardates . . . . .	303
BY ROY P. LANCASTER AND DONALD L. MCMANIS	
Selection of NaCl solutions by sodium-deprived Mongolian gerbils in Richter-type drinking tests . . . . .	315
BY JOSEPH W. CULLEN AND ARTHUR E. HARRIMAN	

A comparison of response rates in response-terminated and time-terminated experimental sessions . . . . .	323
By ALAN C. REPP, SAMUEL M. DEITZ, AND WILLIAM D. WOLKING	
Selective exposure: An addendum . . . . .	329
By LEONARD L. ROSENBAUM AND ELLIOTT MCGINNIES	
Language meaning (gender shaping) among blind and sighted students . .	333
By JOHN G. CULL AND RICHARD E. HARDY	
Effects of genetic counseling on parental self-concepts . . . . .	335
By MARY ANN ANTLEY, RAY M. ANTLEY, AND LAWRENCE C. HARTLAGE	









# ABBREVIATIONS OF WORDS USED IN THE TITLES OF JOURNALS

(One-word titles are never abbreviated. The word "the" is not used, nor its equivalent in any other language. The word "of" or its equivalent in other languages is used only to discriminate what would otherwise be identical titles in different languages. The word "and" is always used, but indicated by "&" in the Roman alphabet. Only English words are indicated here, but the corresponding words in other languages should receive a corresponding abbreviation. All abbreviations and all one-word titles should be in italics.)

Abnormal .....	<i>Abn.</i>	Japanese .....	<i>Jap.</i>
Abstracts .....	<i>Abst.</i>	Journal .....	<i>J.</i>
American .....	<i>Amer.</i>	Mathematical .....	<i>Math.</i>
Anatomy .....	<i>Anat.</i>	Measurement .....	<i>Meas.</i>
Animal .....	<i>Anim.</i>	Medical .....	<i>Med.</i>
Applied .....	<i>Appl.</i>	Mental .....	<i>Ment.</i>
Archives .....	<i>Arch.</i>	Monographs .....	<i>Monog.</i>
Association .....	<i>Assoc.</i>	Neurology .....	<i>Neurol.</i>
Attitude .....	<i>Attit.</i>	Opinion .....	<i>Opin.</i>
Australian .....	<i>Aust.</i>	Orthopsychiatry .....	<i>Orthopsychiat.</i>
Behavior .....	<i>Behav.</i>	Personality .....	<i>Personal.</i>
British .....	<i>Brit.</i>	Personnel .....	<i>Person.</i>
Bulletin .....	<i>Bull.</i>	Philosophy .....	<i>Philos.</i>
Bureau .....	<i>Bur.</i>	Physics .....	<i>Phys.</i>
Canadian .....	<i>Can.</i>	Physiology .....	<i>Physiol.</i>
Character .....	<i>Charac.</i>	Proceedings .....	<i>Proc.</i>
Children .....	<i>Child.</i>	Psychiatry .....	<i>Psychiat.</i>
Chinese .....	<i>Chin.</i>	Psychoanalysis .....	<i>Psychoanal.</i>
Clinical .....	<i>Clin.</i>	Psychology .....	<i>Psychol.</i>
College .....	<i>Coll.</i>	Psychosomatic .....	<i>Psychosomat.</i>
Comparative .....	<i>Comp.</i>	Quarterly .....	<i>Quart.</i>
Consulting .....	<i>Consult.</i>	Religious .....	<i>Relig.</i>
Contributions .....	<i>Contrib.</i>	Research .....	<i>Res.</i>
Development .....	<i>Devel.</i>	Review .....	<i>Rev.</i>
Educational .....	<i>Educ.</i>	School .....	<i>Sch.</i>
Experimental .....	<i>Exper.</i>	Science .....	<i>Sci.</i>
General .....	<i>Gen.</i>	Social .....	<i>Soc.</i>
Genetic .....	<i>Genet.</i>	Statistics .....	<i>Stat.</i>
Indian .....	<i>Ind.</i>	Studies .....	<i>Stud.</i>
Industrial .....	<i>Indus.</i>	Teacher .....	<i>Teach.</i>
International .....	<i>Internat.</i>	University .....	<i>Univ.</i>
Italian .....	<i>Ital.</i>		

# Preparation of Manuscripts for The Journal Press

21 JUN 1973

## GENERAL INSTRUCTIONS

1. The proper sequence for the parts of your submitted manuscript is as follows: (a) text, (b) references, (c) footnotes, (d) tables, (e) figures, and (f) figure legends. However, monographs start with a table of contents and may have an acknowledgment page before the text and an appendix immediately after the text.
2. Use heavy typewriter paper,  $8\frac{1}{2} \times 11$  inches, double-space *all* lines, and leave margins for editorial work. Do not use onionskin or odd sizes of paper.
3. Submit original typewritten copy and one carbon. Retain second carbon for proofing.
4. Retype any page on which written corrections have been made.
5. Do not begin a sentence with a numeral.
6. An adequate summary serving as the conclusion of the text is good form.
7. Each quotation should indicate the page number of the original source. The original publisher must give permission for lengthy quotations and use of tables or figures.
8. Do not fold your manuscript.
9. Enclose your covering letter with your manuscript. If you are unknown to the Editors, kindly state your credentials.

## FORMAT AND SPECIFIC INSTRUCTIONS

### A. TEXT DIVISIONS

#### I. THE TITLES OF JOURNAL ARTICLES AND THE MAJOR SUBDIVISIONS OF MONOGRAPHS ARE PRINTED IN TEN-POINT CAPS CENTERED ON THE PAGE

##### A. THE NEXT SUBDIVISION TITLE IS PRINTED IN CAPS AND SMALL CAPS CENTERED ON THE PAGE

1. *Then italics, with Principal Words, Upper and Lower Case, Centered on the Page*
- a. *Then italics, upper and lower case, 1-em run-in side head.*
  - (1). *Then italics, upper and lower case, 2-em run-in side head.*
  - (a). *Then italics, upper and lower case, 3-em run-in side head.*

[Further subdivision should be merged into the text without marginal indentation, and should be numbered with small letters.]

### B. REFERENCES

References should be arranged in alphabetical order by author, numbered and referred to in the text by number (2).

The proper form of a book reference is as follows:

2. Doe, J. The Preparation of Manuscripts. New York: Holt, 1963. Pp. 400-418.

The proper form of a journal reference is as follows:

2. Doe, J. The preparation of manuscripts. *J. Gen. Psychol.*, 1963, 68, 450-462.

If in the text it is desirable to refer to a page, thus (2, p. 45).

### C. FOOTNOTES

Use as few as possible and number consecutively in the text thus.<sup>1</sup>

<sup>1</sup> Footnote (on the separate footnote page).

### D. TABLES

Tables should always be typed on separate sheets of paper and should be Arabic numbered (Table 2). Construction should be such that vertical lines are unnecessary. Always indicate in the text the approximate insertion place for a table.

### E. FIGURES

Figures should be submitted as glossy prints of the approximate size for final reproduction (ordinarily  $4-4\frac{1}{2}$  inches in width). Lettering and lines should be sharp and clean. Figures should be Arabic numbered (Figure 2) in the text, with the approximate insertion point indicated. Each figure requires a legend, but all legends should be submitted double-spaced on a separate Figure Legends page.

\$39.00 per annum  
\$18.00 per volume  
Single Numbers \$9.00

BIMONTHLY  
Three volumes per year  
Immediate publication

May, 1973  
Volume 84, First Half

Founded in 1933 by Carl Murchison

# The Journal of PSYCHOLOGY

The General Field of Psychology

POWELL MURCHISON  
Managing Editor

HELEN M. DAVIS  
Associate Editor

## Editorial Board

JOHN A. ANTOINETTI  
Life Insurance Agency  
Management Association

S. HOWARD BARTLEY  
Michigan State University

ARTHUR G. BILLS  
University of Cincinnati

E. A. BOTT  
University of Toronto

LEONARD CARMICHAEL  
National Geographic Society

IRVIN L. CHILD  
Yale University

HERBERT S. CONRAD  
U. S. Office of Education

JOHN F. DASHIELL  
Florida State University

LEONARD W. DOOR  
Yale University

If this space should be unstamped, this is the regular library edition.  
But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not  
be resold for a period of five years from date of publication.

FRANK A. GELDARD  
Princeton University

ERNEST R. HILGARD  
Stanford University

JOHN E. HORROCKS  
Ohio State University

DAVID HOTHERSALL  
Ohio State University

HEINRICH KLÜVER  
University of Chicago

NORMAN L. MUNN  
University of Adelaide,  
Australia

WARREN H. TEICHNER  
New Mexico State  
University

LEE EDWARD TRAVIS  
Fuller Theological  
Seminary

JOSEPH ZUBIN  
Psychiatric Institute,  
Columbia University

Copyright © 1973 by The Journal Press  
Provincetown, Massachusetts, U. S. A.

All rights reserved

Second-class postage paid at Provincetown, Mass.  
and additional mailing offices

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

# Back Issues

The Journal Press maintains a complete inventory of back issues for all five journals by reprinting many of them each year. Early volumes frequently consist of multiple parts, and the cost of maintaining these volumes is naturally much higher than that for two-part volumes. So our old policy of a flat price per volume cannot be continued. The following rates are in effect: single numbers, \$9.00; two-part volumes, \$18.00; three-part volumes, \$27.00; four-part and over-four-part volumes, \$36.00. The only exceptions (special issues) are starred below.

	VOLUMES	DATES	VOLUME
<i>Journal of Genetic Psychology</i> (Pedagogical Seminary)			
	1-4	1891-April 1897	\$27.00
	5-37	July 1897-1930	36.00
	38*	1930	18.00
	39	1931	36.00
	40 and later	1932-	18.00
<i>Genetic Psychology Monographs</i>			
	1-22	1926-1940	36.00
	23 and later	1941-	18.00
<i>Journal of General Psychology</i>			
	1-3	1928-1930	36.00
	4*	1930	18.00
	5	1931	36.00
	6 and later	1932-	18.00
<i>Journal of Social Psychology</i>			
	1-10	1930-1939	36.00
	11 and later	1940-	18.00
<i>Journal of Psychology</i>			
	1 and later	1936-	18.00

A mailing charge of \$0.30 per volume is made for foreign deliveries only. Domestic mailings are prepaid. All remittances should be by check or money order payable in U.S. funds on a United States bank. Please do not send UNESCO coupons. Advance payment is required from all individuals and from foreign purchasers. Microfilm versions and special bindings are not available.

THE JOURNAL PRESS  
2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.



*Notice: All copyrighted material reprinted from The Journal Press publications should be given an appropriate acknowledgment. No permission is necessary for brief quotations (less than 300-word total or less than 10 percent of length, whichever is least, for any one article). Permission is required for reprinting tables, figures, entire articles, or substantial portions thereof; and such permission will be contingent upon agreement by the original author or authors. Estate permission may be required. Since no fee is charged for a copyright license, the courtesy of a stamped, self-addressed envelope will be appreciated.*

# The Journal of Psychology

Founded in 1935 by Carl Murchison

The Journal of Psychology was founded at the request of many psychologists throughout the world for the purpose of establishing a medium in which publication shall be immediate and in which other publication experiments may be carried out.

(1). The subscription price is \$39.00 per annum, \$18.00 per volume, \$9.00 per issue (three volumes per year).

(2). The format of each volume is such that each subscriber may tear down the assembled volume into its component articles.

(3). Manuscripts may be sent to any member of the Editorial Board or may be sent directly to The Journal Press. The lag in publication should not be more than a few weeks.

(4). The author will receive 200 reprints of his paper, and subscribers will receive the accumulated publications bound into half-volume units. The second half of each volume will contain the title page and table of contents for that volume. Each half-volume unit will be mailed to the subscribers at its stated date.

(5). Each author will be charged the manufacturing costs of any zincs, halftone cuts, tables, or equations. Excessive use of special fonts (references, footnotes, quotations, appendices) will also be charged.

(6). The publication charges are as follows:

1 to 4 pages	\$40.00
5 to 8 "	80.00
9 to 12 "	120.00
13 to 16 "	160.00
17 to 20 "	200.00
21 to 24 "	240.00
25 to 28 "	280.00
29 to 32 "	320.00
33 to 36 "	360.00
37 to 40 "	400.00

These charges are subject to adjustment as justified.

(7). The accumulated units will be mailed to subscribers bimonthly: January, March, May, July, September, and November. The units may vary in size to accommodate publication needs.

All back issues are available. See announcement on inside front cover.

THE JOURNAL PRESS

2 Commercial Street

Provincetown, Massachusetts 02657

U. S. A.

# Journal of Genetic Psychology

Founded by G. Stanley Hall in 1891; reorganized by Carl Murchison in 1925

Quarterly, beginning with March. Devoted to developmental and clinical psychology, with briefly reported Replications and Refinements and occasional book reviews. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$26.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

---

# Journal of General Psychology

Founded in 1927 by E. B. Titchener and Carl Murchison

Quarterly, beginning with January. Devoted to experimental, physiological, and comparative psychology, with briefly reported Replications, Refinements, and Comments. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$26.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

---

# Journal of Social Psychology

Founded in 1929 by John Dewey and Carl Murchison

Bimonthly, beginning with February. Devoted to studies of persons in group settings and of culture and personality; special attention to cross-cultural articles and notes, and to briefly reported Replications and Refinements. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Three volumes annually. Per annum, \$39.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

---

*Each author will receive 100 unbound reprints gratis. Additional reprints may be purchased at a nominal cost when the author returns page-proof.*

THE JOURNAL PRESS  
2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.

# Club Rate for Five International Psychological Journals

	PER ANNUM
THE JOURNAL OF GENETIC PSYCHOLOGY . . . . . (Quarterly, two volumes per year)	\$26.00
GENETIC PSYCHOLOGY MONOGRAPHS . . . . . (Quarterly, two volumes per year)	\$26.00
THE JOURNAL OF GENERAL PSYCHOLOGY . . . . . (Quarterly, two volumes per year)	\$26.00
THE JOURNAL OF SOCIAL PSYCHOLOGY . . . . . (Bimonthly, three volumes per year)	\$39.00
THE JOURNAL OF PSYCHOLOGY . . . . . (Bimonthly, three volumes per year)	\$39.00
	<b>TOTAL \$156.00</b>
TEACHERS AND STUDENTS EDITION, <i>Restricted Sale</i>	\$50.00

The *Teachers and Students Edition* is open to any individual psychologist, teacher, student, or physician but requires an agreement that the edition will not be offered for sale till after five years from publication and an agreement that the edition will not be substituted for an institutional subscription. Write for our Teachers and Students contract card.

THE JOURNAL PRESS  
2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.

## Genetic Psychology Monographs

Founded in 1925 by Carl Murchison

Quarterly, beginning with February. Devoted to developmental and clinical psychology. Each number one or more complete researches. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$26.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

**NOTICE.** Effective immediately we offer prior publication for certain selected monographs. Charges will be based on the announced rates for *The Journal of Psychology*. Regular monographs will carry no page charge for the first 50 pages, but for all pages above 50 there will be a charge of \$10.00 per page. Manufacturing costs of any zincs, halftone cuts, tables, or equations will be charged to authors, as will excessive use of special fonts. Two hundred reprints of each monograph will be supplied free of charge.

The listing of monographs for 1926 through 1963 will be printed in the first half of each volume of all five journals. Monographs published after 1963 will be listed in the second half of each volume.

# Genetic Psychology Monographs (continued)

VOLUME 1-1926

1. Performance tests for children of pre-school age—R. STUTSMAN
2. An experimental study of the eidetic type—H. KLÜVER
- 3 & 4. A study of natio-racial mental differences—N. D. M. HIRSCH
5. A psychological study of juvenile delinquency by group methods—J. W. BRIDGES AND K. M. B. BRIDGES
6. The influence of puberty praecox upon mental growth—A. GESSLL

VOLUME 2-1927

- 1 & 2. The mind of a gorilla—R. M. YERKES
3. The role of eye-muscles and mouth-muscles in the expression of the emotions—K. DUNLAP
4. Family similarities in mental-test abilities—R. R. WILLOUGHBY
5. Coordination in the locomotion of infants—L. H. BURNSIDE
6. The mind of a gorilla: Part II. Mental development—R. M. YERKES

VOLUME 3-January-June, 1928

1. An experimental study of the olfactory sensitivity of the white rat—J. R. LIGGETT
2. A photographic study of eye movements in reading formulae—M. A. TINKER
3. An experimental study of the East Kentucky mountaineers—N. D. M. HIRSCH
4. Responses of foetal guinea pigs prematurely delivered—G. T. AVERY
5. Objective differentiation between three groups in education (teachers, research workers, and administrators)—M. B. JENSEN
6. The effect of segregation on the sex behavior of the white rat as measured by the obstruction method—M. JENSEN

VOLUME 4-July-December, 1928

1. Observation and training of fundamental habits in young children—E. A. BOTT, W. E. BLATZ, N. CHANT, AND H. BOTT
- 2 & 3. Determination of a content of the course in literature of a suitable difficulty for junior and senior high school students—M. C. BURCH
- 4 & 5. Methods for diagnosis and treatment of cases of reading disability—M. MONROE
6. The relative effectiveness of lecture and individual reading as methods of college teaching—E. B. GREENE

VOLUME 5-January-June, 1929

1. The age factor in animal learning: I. Rats in the problem box and the maze—C. P. STONE
2. The effect of delayed incentive on the hunger drive in the white rat—E. L. HAMILTON
3. Which hand is the eye of the blind?—J. M. SMITH
4. The effect of attitude on free word association-time—A. G. EKDAHL
5. The localization of tactual space: A study of average and constant errors under different types of localization—L. E. COLE
6. The effects of gonadectomy, vasotomy, and injections of placental and orchic extracts on the sex behavior of the white rat—H. W. NISSEN

VOLUME 6-July-December, 1929

1. Learning and growth in identical infant twins: An experimental study by the method of co-twin control—A. GESSLL AND H. THOMPSON
2. The age factor in animal learning: II. Rats on a multiple light discrimination box and a difficult maze—C. P. STONE
3. The acquisition and interference of motor habits in young children—E. MCGINNIS
4. A vocational and socio-educational survey of graduates and non-graduates of small high schools of New England—A. D. MUELLER
- 5 & 6. A study of the smiling and laughing of infants in the first year of life—R. W. WASHBURN

VOLUME 7-January-June, 1930

1. Tensions and emotional factors in reaction—E. DUFFY
2. Teacher influence on class achievement: A study of the relationship of estimated teaching ability to pupil achievement in reading and arithmetic—H. R. TAYLOR
- 3 & 4. A study of the effect of inverted retinal stimulation upon spatially coordinated behavior—P. H. EWERT
5. A study of the mental development of children with lesion in the central nervous system—E. E. LOND
6. An experimental study upon three hundred school children over a six-year period—N. D. M. HIRSCH

VOLUME 8-July-December, 1930

1. The amount and nature of activities of newborn infants under constant external stimulating conditions during the first ten days of life—O. C. LEWIN
2. Race and social differences in performance tests—S. D. PORTKUS, et al.
3. Language and growth: The relative efficacy of early and deferred vocabulary training, studied by the method of co-twin control—L. C. STRAYER
4. Eye-movements and optic nystagmus in early infancy—J. M. MCGINNIS
- 5 & 6. Reactions of kindergarten, first-, and second-grade children to constructive play materials—L. FARWELL

VOLUME 9-January-June, 1931

- 1 & 2. The status of the first-born with special reference to intelligence—H. H. HSIAO
- 3 & 4. An experimental study of bright, average, and dull children at the four-year mental level—H. P. DAVIDSON
5. An historical, critical and experimental study of the Scashore-Kwalwasser test battery—P. R. FARNSWORTH
6. A comparison of difficulty and improvement in the learning of bright and dull children in reproducing a descriptive selection—F. T. WILSON

VOLUME 10-July-December, 1931

1. A comparative study of a group of southern white and negro infants—M. B. MCGRAW
- 2 & 3. An experimental study of prehension in infants by means of systematic cinema records—H. M. HALVERSON
4. The limits of learning ability in kittens—A. M. SHUEY
- 5 & 6. The effect of habit interference upon performance in maze learning—O. W. ALM

VOLUME 11-January-June, 1932

1. General factors in transfer of training in the white rat—T. A. JACKSON
2. The effect of color on visual apprehension and perception—M. A. TINKER
3. The reliability and validity of maze experiments with white rats—R. LEEPER
4. A critical study of two lists of best books for children—F. K. SHUTTLEWORTH
- 5 & 6. Measuring human energy cost in industry: A general guide to the literature—R. M. PAGE

VOLUME 12-July-December, 1932

1. Family resemblances in verbal and numerical abilities—H. D. CARTER
2. The development of fine prehension in infancy—B. M. CASTNER
- 3 & 4. The growth of adaptive behavior in infants: An experimental study at seven age levels—H. M. RICHARDSON
- 5 & 6. Differential reactions to taste and temperature stimuli in newborn infants—K. JENSEN

VOLUME 13-January-June, 1933

1. A critique of sublimation in males: A study of forty superior single men—W. S. TAYLOR
2. A study of the nature, measurement, and determination of hand preference—H. L. KOCH, et al.
3. The growth and decline of intelligence: A study of a homogeneous group between the ages of ten and sixty—H. E. JONES AND H. S. CONRAD
4. The relation between the complexity of the habit to be acquired and the form of the learning curve in young children—M. L. MATTHEWSON
5. Eating habits in relation to personality development of two- and three-year-old children: A study of sixty-nine children in two nursery schools—A. A. ELIOT
6. Coordinating mechanisms of the spinal cord—O. C. INGBERSSEN



# Genetic Psychology Monographs (continued)

VOLUME 14—July-December, 1933

- Mental growth during the first three years: A developmental study of sixty-one children by repeated tests—N. BAYLEY
- A study of triplets: including theories of their possible genetic relationships—F. N. ANDERSON AND N. V. SCHEIDEMANN
- The objective measurement of emotional reactions—H. V. GASKILL
- Development of behavior in the fetal cat—J. D. CORONIOS
- A study of certain language developments of children in grades four to twelve, inclusive—L. L. LABRANT
- The effect of early and delayed practice on memory and motor performances studied by the method of co-twin control—J. R. HILGANO

VOLUME 15—January-June, 1934

- Studies in the psychology of tone and music—P. R. FARNSWORTH
- Motor learning of children in equilibrium in relation to nutrition—E. L. BEERE
- Discrimination limits of pattern and size in the goldfish *Carassius auratus*—J. B. ROWLEY
- Limits of learning ability in the white rat and the guinea pig—B. F. RIES
- The limits of learning ability in rhesus monkeys—H. A. FIELD

VOLUME 16—July-December, 1934

- A statistical study of ratings on the California Behavior Inventory for Nursery-School Children—H. S. CONRAD
- An eye-movement study of objective examination questions—A. FRANDSEN
- An experimental study of constitutional types—O. KLINGER, S. E. ASCH, AND H. BLOCK
- The development of a battery of objective group tests of manual laterality, with the results of their application to 1900 children—W. N. DUKORF
- An experimental study in the prenatal guinea-pig of the origin and development of reflexes and patterns of behavior in relation to the stimulation of specific receptor areas during the period of active fetal life—L. CARMICHAEL

VOLUME 17—January-December, 1935

- Organization of behavior in the albino rat—R. L. THORNDIKE
- Brightness discrimination in the rhesus monkey—M. P. CHAWFORD
- The limits of learning ability in cebus monkeys—A. M. KOCU
- Nature-nurture and intelligence—A. M. LEAHY
- On intelligence of epileptic children—E. B. SULLIVAN AND L. GANAGAN
- A study of the play of children of preschool age by an unobserved observer—D. L. COCKRELL

VOLUME 18—January-December, 1936

- Sex differences in variational tendency—Q. MCNEAMAR AND L. M. TERMAN
- The process of learning to dress among nursery-school children—C. B. KEY, M. R. WHITE, M. P. HONZIK, A. B. HINNEY, AND D. ERWIN
- A study of the present social status of a group of adults, who, when they were in elementary schools, were classified as mentally deficient—W. R. BALLER
- The influence of specific experience upon mental organizations—A. ANASTASI
- Studies in aggressiveness—L. BENDER, S. KEISER, AND P. SCHILDER

VOLUME 19—January-December, 1937

- Psychological bases of self mutilation—C. DABROWSKI
- Intelligence, temperament and secondary sex characteristics: A study of the relationship between psychological and physical measures of masculinity—H. GULKINSON
- A psychological study of forty unmarried mothers—R. D. NOTTINGHAM
- Behavior problems in the children of psychotic and criminal parents—L. BENDER
- Domination and integration in the social behavior of young children in an experimental play situation—H. H. ANDERSON
- The sequential patterning of prone progression in the human infant—L. B. AMES

VOLUME 20—January-December, 1938

- The relationship between characteristics of personality and physique in adolescents—P. S. DE Q. CABOT
- Behavior problems of elementary school children: A descriptive and comparative study—I. Y. MASTEN
- Graphic representation of a man by four-year-old children in nine prescribed drawing situations—P. F. GRIDLEY
- Differences between two groups of adult criminals—R. S. TOLMAN
- A comparative study by means of the Rorschach method of personality development in twenty pairs of identical twins—E. TROPE
- Individual differences in the facial expressive behavior of preschool children: A study by the time-sampling method—C. SWAN

VOLUME 21—January-December, 1939

- An experimental analysis of "level of aspiration"—R. GOULD
- Some light on the problem of bilingualism as found from a study of the progress in mastery of English among preschool children of non-American ancestry in Hawaii—M. E. SMITH
- Domination and social integration in the behavior of kindergarten children and teachers—H. H. ANDERSON
- The capacity of the rhesus and cebus monkey and the gibbon to acquire differential response to complex visual stimuli—W. E. GALT
- The social-sex development of children—E. H. CAMPBELL

VOLUME 22—January-December, 1940

- Measuring human relations: An introduction to the study of the interaction of individuals—E. D. CHAPPEL
- Aggressive behavior in young children and children's attitudes toward aggression—M. D. FITZ
- Student attitudes toward religion—E. NELSON
- The prediction of the outcome on furlough of dementia praecox patients—J. S. JACOB
- Significant characteristics of preschool children as located in the Conrad inventory—K. H. READ
- Learning by children at noon-meal in a nursery school: Ten "good" eaters and ten "poor" eaters—J. B. McCAY, E. B. WAINIC, AND P. J. KAUSZ
- Studies in the interpretation of play: I. Clinical observation of play disruption in young children—E. H. ERIKSON

VOLUME 23—January-June, 1941

- An analysis of certain variables in a developmental study of language—F. M. YOUNG
- Infant development under conditions of restricted practice and of minimum social stimulation—W. DENNIS
- An analysis of the mental factors of various age groups from nine to sixty—B. BALINSKY
- Factors influencing performance on group and individual tests of intelligence: I. Rate of work—M. W. BENNETT
- Individual differences in appetitive reaction: A study of the response of preschool children to pictures—E. W. AMER

VOLUME 24—July-December, 1941

- Twins T and C from infancy to adolescence: A biogenetic study of individual differences by the method of co-twin control—A. GESELL AND H. THOMPSON
- Anger out-biting: Its incipency, incidence, and amelioration—A. L. BULLIC
- An experimental study of the factors of maturation and practice in the behavioral development of the embryo of the frog, *Rana pipiens*—A. FROMME
- Measurement of the size of general English vocabulary through the elementary grades and high school—M. K. SMITH
- Genotypes in the field of musical eminence—P. R. FARNSWORTH

VOLUME 25—January-June, 1942

- A study of factors determining family size in a selected professional group—J. C. FLANAGAN
- Genetic study of geometrical-optical illusions—A. WALTERS
- Interpretation of behavior-ratings in terms of favorable and unfavorable deviations: A study of scores from the Read-Conrad Behavior Inventory—K. H. READ AND H. S. CONRAD
- Are there any innate behavior tendencies?—J. B. SCHOOLLAND
- An investigation of the intelligibility of the speech of the deaf—C. V. HUDGINS AND F. C. NUMBERS



# Genetic Psychology Monographs (continued)

VOLUME 26—July-December, 1942

1. The critical frequency limen for visual flicker in children between the ages of 6 and 18—V. L. MULLER  
Some factors determining handedness in the white rat—K. L. WENTWORTH
2. Motivation and behavior—E. FRENKEL-BRUNSWIK

VOLUME 27—January-June, 1943

1. Comparison of children's personality traits, attitudes, and intelligence with parental occupation—N. R. MACEY
2. A comparative study of mental functioning patterns of problem and non-problem children seven, eight, and nine years of age—M. L. PIGNATELLI

VOLUME 28—July-December, 1943

1. Separation anxiety in young children: A study of hospital cases—H. EDELSTON
2. Correlates of vocational preferences—W. A. BRADLEY, JR.

VOLUME 29—January-June, 1944

1. Mental changes after bilateral prefrontal lobotomy—S. D. PORTEUS AND R. D. KEPNER
2. A twin-controlled experiment on the learning of auxiliary languages—B. PRICE, W. J. KOSTER, AND W. M. TAYLOR

VOLUME 30—July-December, 1944

1. A method of administering and evaluating the thematic appreciation test in group situations—R. M. CLARK
2. A study of anxiety reactions in young children by means of a projective technique—R. TEMPLE AND E. W. AMEN

VOLUME 31—January-June, 1945

1. The evolution of intelligent behavior in rhesus monkeys—B. WEINSTEIN
2. Perceptual behavior of brain-injured, mentally defective children: An experimental study by means of the Rorschach technique—H. WERNER

VOLUME 32—July-December, 1945

1. A clinical study of sentiments: I.—H. A. MURRAY AND C. D. MORGAN
2. A clinical study of sentiments: II.—H. A. MURRAY AND C. D. MORGAN

VOLUME 33—January-June, 1946

1. Interpretation of spontaneous drawings and paintings—T. S. WAERNER  
Preferences for sex symbols and their personality correlates—K. FRANCK
2. Outstanding traits: In a selected college group, with some reference to career interests and war records—F. L. WOODS AND W. L. WOODS

VOLUME 34—July-December, 1946

1. The relation of emotional adjustment to intellectual function—J. L. DESPERT AND H. O. PIERCE  
The smiling response: A contribution to the ontogenesis of social relations—R. A. SPITZ
2. Finger-painting and personality diagnosis—P. J. NAPOLI

VOLUME 35—January-June, 1947

1. The thematic apperception technique in the study of culture-personality relations—W. E. HENRY
2. A continuation study of anxiety reactions in young children by means of a projective technique—M. DONKIN E. W. AMEN  
A study of the vocational interest trends of secondary school and college women—A. M. CAWLEY

VOLUME 36—July-December, 1947

1. Make test validation and psychosurgery—S. D. PORTEUS AND H. N. PETERS
2. The diagnostic implications of Rorschach's test in case studies of mental defectives—I. JOLLES

VOLUME 37—January-June, 1948

1. The radio day time serial: A symbolic analysis—W. L. WARNER AND W. E. HENRY  
The relation of personality characteristics and response to verbal approval in a learning task—G. L. GRACE
2. The mechanism of vision: XVIII. Effects of destroying the visual "associative areas" of the monkey—K. S. LEE  
A study of the relationship between handwriting and personality variables—P. CASTELNUOVA-TEDESCO

VOLUME 38—July-December, 1948

1. Modern language learning: The intensive course as sponsored by the United States Army, and implications for undergraduate course of study—M. LIND  
Conflict: A study of some interactions between appetite and aversion in the white rat—M. A. TOLCOTT
2. Schizophrenia and the MAPS test: A study of certain formal psycho-social aspects of fantasy production in schizophrenia as revealed by performance on the Make a Picture Story (MAPS) Test—E. S. SHNEIDMAN  
A study of the transmission of authority patterns in the family—H. L. INGERSOLL

VOLUME 39—January-June, 1949

1. A study of the psychoanalytic theory of psychosexual development—G. S. BLUM  
The assessment of parental attitudes in relation to child adjustment—E. J. SHOEN, JR.
2. Qualitative differences in the vocabulary responses of normals and abnormals—H. FEIPEL  
The relative effectiveness of motion and still pictures as stimuli for eliciting fantasy stories about adolescent relationships—P. E. EISNER  
The organization of hereditary maze-brightness and maze-dullness—L. V. SEARLE

VOLUME 40—July-December, 1949

1. An experimental study of what young school children expect from their teachers—R. BIBER AND C. LEWIS  
A study of the relative effects of age and of test difficulty upon factor patterns—H. A. CLINTS
2. A projective experiment using incomplete stories with multiple choice endings—J. K. SEATON  
Effects of sex role and social status on the early adolescent personality—E. MILNER  
Social perceptions and attitudes of children—M. RADKE, H. TRACER, AND H. DAVIS

VOLUME 41—January-June, 1950

1. Some psychological and educational aspects of pediatric practice: A study of well-baby clinics—L. H. BLUM  
One-trial learning in the domestic rat—B. B. HUDSON  
An introduction to the principles of scientific psychoanalysis—A. ELLIS
2. Awareness of racial differences by preschool children in Hawaii—D. V. SPRINGER  
Age trends in children's evaluation of teacher-approved and teacher-disapproved behavior—S. L. WITRYOL  
The relationship between level of vocational aspiration and certain personal data: A study of some traits and influences bearing on the prestige level of vocational choice—J. STUBBINS

VOLUME 42—July-December, 1950

1. Personality patterns of suicidal mental hospital patients—N. L. FARBROW  
Sex-role identification in young children in two diverse social groups—M. RABAN
2. A study of the influence of the social field on individual behavior: As revealed in the expression of hostility by warmth by neurotics and paranoid schizophrenics in discussion group situations—D. SHAPIRO  
An experimental study of avoidance—R. F. HOFFERLINE

VOLUME 43—January-June, 1951

1. A study of copying ability in children—E. A. TOWNSEND  
Prestige motivation of gifted children—D. P. AUSUBEL
2. A psychological study of physical scientists—A. ROSE

VOLUME 44—July-December, 1951

1. The organization of hostility controls in various personality structures—S. FISHER AND E. HINDS  
Children and radio: A study of listeners and non-listeners to various types of radio programs in terms of personality ability, attitude, and behavior measures—E. A. RICCIUTI
2. Quantitative expression in young children—W. E. MARTIN  
The use of magnetic devices in the collection and analysis of the preverbal utterances of an infant—A. W. LIND

VOLUME 45—January-June, 1952

1. Japanese-American personality and acculturation—W. CAUDILL
2. A statistical study of the Freudian theory of levels of psychosexual development—C. A. BARNES  
Personality characteristics of selected disability groups—D. N. WINKER

# Genetic Psychology Monographs (continued)

VOLUME 46—July-December, 1952

- The relationship of social status, intelligence, and sex of ten- and eleven-year-old children to an awareness of poverty—F. J. ESTVAN  
 A critical study of the castration and Oedipus complexes—S. M. FRIEDMAN  
 The relationship between projective test scoring categories and activity preferences—M. M. SCHWARTZ  
 A comparison of formal and content factors in the diagnostic testing of schizophrenia—M. SHERMAN

VOLUME 47—January-June, 1953

- Ability and accomplishment of persons earlier judged mentally deficient—D. C. CHARLES  
 Variations in the consistency of the behavioral meaning of personality test scores—M. KORNREICH  
 Some child-rearing antecedents of aggression and dependency in young children—R. R. SEARS, et al.  
 Symptom correlates for descriptive diagnosis—J. R. WITTENBORN, et al.

VOLUME 48—July-December, 1953

- Age and mental abilities: A longitudinal study—W. A. OWENS, JR.  
 The development of a personality questionnaire for drinkers—P. J. HAMPTON  
 Personality and physical disease: A test of the Dunbar hypothesis applied to diabetes mellitus and rheumatic fever—D. H. CROWELL  
 Socio-economic contrasts in children's peer culture prestige values—B. POPE  
 A critical review of the stability of social acceptability scores obtained with the partial-rank-order and the paired-comparison scales—S. A. WITRYOL AND G. G. THOMPSON  
 A study of the effects of color on Rorschach responses—G. G. BRODY

VOLUME 49—January-June, 1954

- Factors underlying major reading disabilities at the college level—J. A. HOLMES  
 Parent behavior toward first and second children—J. K. LASKO  
 Social status and intelligence: An experimental study of certain cultural determinants of measured intelligence—E. A. HAGGARD  
 Certain determinants and correlates of authoritarianism—S. SIEGEL  
 Personalities in faces: I. An experiment in social perceiving—P. F. SECORD, W. F. DUKES, AND W. BEVAN

VOLUME 50—July-December, 1954

- A study of the relationship between play patterns and anxiety in young children—E. W. AMEN AND N. RENISON  
 Operational exploration of the conceptual self system and of the interaction between frames of reference—M. EDELSON AND A. E. JONES  
 Problem solving: A statistical description of some relationships between organismic factors and selected response measures—N. A. FATTU, E. KAPOS, AND E. V. MECH  
 The relation of cortical potentials to perceptual functions—C. CHYATTE  
 The impact for clinical psychology of the use of tests derived from theories about infantile sexuality and adult character—D. W. MILES  
 Measuring personality in developmental terms: The Personal Preference Scale—M. H. KROUT AND J. K. TABIN

VOLUME 51—January-June, 1955

- Some relations between techniques of feeding and training during infancy and certain behavior in childhood—A. BRENNSTEIN  
 The expression of personality in drawings and paintings—L. H. STEWART  
 Negative stereotypes concerning Americans among American-born children receiving various types of minority-group education—J. A. FISHMAN  
 The Lincoln-Oseretsky Motor Development Scale—W. SLOAN

VOLUME 52—July-December, 1955

- Some personality correlates of sex, sibling position, and sex of sibling among five- and six-year-old children—H. L. KOCIR  
 A quantitative Rorschach assessment of maladjustment and rigidity in acculturating Japanese Americans—C. DeVos  
 Measurement of authoritarianism and its relation to teachers' classroom behavior—H. M. MCGEE  
 The formal aspects of schizophrenic verbal communication—B. MIRIN  
 A study in an aspect of concept formation, with subnormal, average, and superior adolescents—H. N. HOFFMAN  
 Traumatic avoidance learning: Acquisition and extinction in dogs deprived of normal peripheral autonomic function—L. C. WYNNE AND R. L. SOLOMON

VOLUME 53—January-June, 1956

- As the psychiatric aide sees his work and problems—F. L. WELLS, M. GREENBLATT, AND R. W. HYDE  
 An investigation of avoidance, anxiety, and escape behavior in human subjects as measured by action potentials in muscle—J. D. BROTHURS  
 Spread of effect: A critical review—M. H. MARX  
 Grief, fantasy, and schizophrenia: A study of the adaptive processes—O. J. B. KERNER  
 The attitude structure of the individual: A Q-study of the educational attitudes of professors and laymen—F. N. KERLINGER

VOLUME 54—July-December, 1956

- A study of personality differences between middle and lower class adolescents: The Scondi Test in culture-personality research—L. RAINWATER  
 The assessment of parental identification—S. W. GRAY AND R. KLAUS  
 The influence of social context on impulse and control tendencies in preadolescents—G. H. ZUK  
 Under-mindedness versus tough-mindedness in psychology: A reexamination—H. WINTHROP  
 A method for the comparison of groups: A study in thematic apperception—L. C. SCHAW AND W. E. HENRY

VOLUME 55—January-June, 1957

- Academic performance and personality adjustments of highly intelligent college students—B. M. HORRALL  
 The use of the Vineland Social Maturity Scale in the planning of an educational program for non-institutionalized low-grade mentally deficient children—M. G. GOTTSCHEN  
 The structure and origin of the anal character—H. BELOFF  
 The expression of adolescents' interests—M. AMATORA  
 The role of mass media and the effect of aggressive film content upon children's aggressive responses and identification choices—R. S. ALBERT

Interest in persons as an aspect of sex difference in the early years—E. W. GOODENOUGH

VOLUME 56—July-December, 1957

- Stable response determinants of perception, thinking, and learning: A study based on the analysis of a single test—M. L. SIMMEL AND S. COUNTS  
 Potentials of age: An exploratory field study—S. L. PRESSEY  
 Verbal communication—L. K. FRANK  
 The use of a filmed puppet show as a group projective technique for children—M. R. HAWORTH  
 The social competence of middle-aged people—R. J. HAVIGHURST

VOLUME 57—January-June, 1958

- Psychological and cultural problems in mental subnormality: A review of research—S. B. SARASON AND T. GLADWIN  
 Developmental aspects of discrimination in relation to adjustment—P. LONDON  
 Muscular tension as a response to psychological stress in rheumatoid arthritis and peptic ulcer—J. A. SOUTHWORTH  
 Irritation: Theory and experiment—R. LAWSON AND M. H. MARX

VOLUME 58—July-December, 1958

- Emotional aspects of political behavior: The woman voter—E. M. BENNETT AND H. M. GOODWIN  
 The accuracy of self estimate: A measure of self-concept reality—R. M. BRANDT  
 Personality factors in social mobility: A study of occupationally mobile businessmen—J. C. ABERGLEN  
 Interaction between authoritarian and nonauthoritarian principals and teachers—P. LAMBERT  
 Personality factors related to the choice-status of ninety eighth grade children in a school society—D. ELKINS  
 The relationships between fantasy and self-report behavior—R. C. CALOGERAS

# Genetic Psychology Monographs (continued)

VOLUME 59—January-June, 1959

1. Freshmen Rorschachs and college performance—G. R. SCHMEIDLER, M. J. NELSON, AND M. BRISTOL  
Perceptual aspects of attributed movement—R. W. COAN  
Men and women: Personality patterns and contrasts—E. M. BENNETT AND L. R. COHEN
2. Personality factors in mothers of cerebral palsied children—G. BOLES  
Separatism and integrationism: A social-psychological analysis of editorial content in New York newspapers of three American minority groups—J. A. FISHMAN AND G. S. FISHMAN  
Self, role, and satisfaction—A. L. BROPHY

VOLUME 60—July-December, 1959

1. The constancy of personality ratings over two decades—R. D. TUDDENHAM  
Mother-son identification and vocational interest—L. H. STEWART  
A study of mother-child relationships in the emotional disorders of children—M. J. ROSENTHAL, M. FINKELSTEIN, E. N. AND R. E. ROBERTSON  
On the trail of the wolf-children—W. F. OGBURN AND N. K. BOSS
2. Measuring the mental health of normal adults—R. F. PECK  
Parental identification in young children—W. EMMERICH  
The problems of Jewish religious youth—B. M. LEVINSON

VOLUME 61—January-June, 1960

1. Identifying socially maladjusted school children—A. A. ADAMS  
Social interaction in an interracial nursery school—H. W. STEVENSON AND N. G. STEVENSON, *et al.*  
Interests of pre-adolescent boys and girls—M. AMATORA  
Educational and sex differences in the organization of abilities in technical and academic students in Colombia, South America—J. F. FRIELLA
2. Some themes in the personalities of German men—L. RAINWATER  
Right-left gradients in body image, body reactivity, and perception—S. FISHER  
Longitudinal survey of child Rorschach responses: Younger subjects two to 10 years—L. B. AMES  
Personality patterns related to occupational roles—M. SIEGELMAN AND R. F. PECK

VOLUME 62—July-December, 1960

1. The reading process: A descriptive, interdisciplinary approach—M. P. LANGMAN  
Reactions to stress as a function of the level of intelligence—S. M. H. ZAHDI  
Rorschach diagnosis by a systematic combining of content, thought process, and determinant scales—P. A. BOWEN, R. TESTIN, AND A. ROBERTS  
Longitudinal survey of child Rorschach responses: Older subjects aged 10 to 16 years—L. B. AMES
2. Twenty years of shock therapy in America, 1937-1956: An annotated bibliography—N. H. PRONKO, R. SITTERLY, AND K. BROWN  
The quantitative analysis of parent behavior toward psychotic children and their siblings—E. M. DONNELLY

VOLUME 63—January-June, 1961

1. Make a sentence test: An approach to objective scoring of sentence completions—E. F. BORGATTA  
Some aspects of the relationship between perception and motility in children—E. P. ROTHMAN  
Some psychophysical aspects of mental illness in children—P. H. BERKOWITZ  
Conceptual thinking in schizophrenic children—G. FRIEDMAN
2. Some maternal influences on children's personality and character—J. C. FINNEY  
A study of concept formation in normals, mental defectives, and brain-damaged adults—R. W. ZARLOW  
An investigation of trait relationships among six-year-old children—S. A. AVAKIAN

VOLUME 64—July-December, 1961

1. Coping patterns of preschool children in response to intelligence test demands—A. MORIARTY  
Studies of individual and paired interactional problem-solving behavior of rats: II. Solitary and social controls—A. BARON AND R. A. LITTMAN
2. Studies of role performance—J. H. MANN  
A psychological assessment of professional actors and related professions—R. TAFT  
A cross-cultural study of menstrual taboos—W. N. STEPHENS  
The biological roots of creativity—H. GUTMAN

VOLUME 65—January-June, 1962

1. A Rorschach study of the development of personality structure in white and Negro children in a Southeastern community—A. C. PRICE  
Measuring personality patterns of women—J. LOEVINGER  
Home interests in early adolescence—M. AMATORA  
Social and emotional adjustment during adolescence as related to the development of psychosomatic illness in adulthood—L. H. STEWART
2. A systematic study of interaction process scores, peer and self-assessments, personality and other variables—E. F. BORGATTA  
Children's conceptions of the content and functions of the human body—E. GILBERT

VOLUME 66—July-December, 1962

1. Ocular-manual laterality and perceptual rotation of literal symbols—L. FLESCHER  
Parental child-rearing attitudes and their correlates in adolescent hostility—S. B. CHOROST  
Some explorations into the nature of anxieties relating to illness—P. R. ROBBINS  
Developmental-interactive aspects of hypnosis: Some illustrative cases—J. R. HILCARD AND E. R. HILCARD
2. Teaching a two-year-old to read: An experiment in early childhood learning—W. FOWLER  
Achievement gains or losses during the academic year and over the summer vacation period: A study of trends in achievement by sex and grade level among students of average intelligence—K. M. PARSLEY, JR., AND M. POWELL

VOLUME 67—January-June, 1963

1. Verbatim dialogue of a mother and child in therapy—C. E. MOUSTAKAS  
A study of self-concept: Delinquent boys' accuracy in selected self-evaluations—W. E. AMOS  
An integrating study of the factor structure of adult attitude-interests—R. B. CATTELL AND J. HORN  
Social factors in mental disorders: A contribution to the etiology of mental illness—W. C. COBLINER
2. Conceptual preferences associated with choice of quantitative and nonquantitative college majors—J. GILBERT  
A study of the concept of the feminine role of 51 middle-class American families—A. STEINMANN

VOLUME 68—July-December, 1963

1. Brain dynamism as reflected in illusions—G. K. YACORZYNSKI  
The nature and measurement of components of motivation—R. B. CATTELL, J. A. RADCLIFFE, AND A. B. SWENEY
2. Sharing in preschool children as a function of amount and type of reinforcement—W. F. FISCHER  
The Gesell Incomplete Man Test as a measure of developmental status—L. B. AMES AND F. L. ILLS  
Psychological and sociological factors in prediction of career patterns of women—M. C. MULVEY  
Functionalism and its historical significance—R. HARRISON



## PUBLICATION IN THE JOURNAL OF PSYCHOLOGY

*The Journal of Psychology* provides a special service to authors and readers. Articles accepted for publication appear in print with the least possible delay. Although editorial standards are the same as those for the other journals, all correspondence and handling are expedited, and the referees may, at their discretion, permit some latitude in length of treatment with respect to tabular material, discussion, and review of the literature.

To save time, authors are requested to exercise unusual care in preparing their papers. Special attention should be paid to the General Instructions on the back cover of each issue and to the announcement on the first preliminary page.



\$39.00 per annum  
\$18.00 per volume  
Single Numbers \$9.00

BIMONTHLY  
Three volumes per year  
Immediate publication

May, 1973  
Volume 84, First Half

(Founded in 1935 by Carl Murchison)

# The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a bilateral contract, and may not be resold for a period of five years from the date of publication.

MAY, 1973

(Manuscripts are printed in the order of final acceptance)

Judgment of relative recency: Developmental trends . . . . .	3
By J. M. VON WRIGHT	
Manipulation of situational favorability in tests of the contingency model . . . . .	13
By L. H. GEYER AND J. W. JULIAN	
The effects of a zero interval on semantic differential rotated factor loadings . . . . .	23
By JOHN H. FORTMAN	
Motivational determinants of family planning clinic attendance . . . . .	33
By CARL M. COCHRANE, CLARK E. VINCENT, C. ALLEN HANEY, AND ROBERT MICHIELUTTE	
Name style and conservatism . . . . .	45
By ROGER BOSHIER	

(OVER)

Copyright © 1973 by The Journal Press  
Provincetown, Massachusetts, U. S. A.

All rights reserved

Second-class postage paid at Provincetown, Mass  
and additional mailing offices

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

Polarity and "accuracy" of ratings and the meaningfulness of personality dimensions . . . . .	55
BY WALTER C. BORMAN AND WILLIAM K. GRAHAM	
Premackian reinforcement of classroom behavior through topic sequencing . . . . .	61
BY JACK C. HARTJE	
Conservation of space in noninstitutionalized old people . . . . .	75
BY DIANE E. PAPALIA, ELIZABETH KENNEDY, AND NANCY SHEEHAN	
Effects of external genital sensory feedback on copulatory behavior of rats . . . . .	81
BY PETER O. PERETTI	
Pronunciability in verbal learning of the deaf . . . . .	89
BY KATHLEEN CHEN	
Self-image disparity and attachment to ethnic subculture . . . . .	97
BY MARK MCCORMICK AND DAVID BALLA	
Self-perception differences among kibbutz and city adults in Israel and Jewish and non-Jewish adults in the United States . . . . .	105
BY LEO A. PIROJNICKOFF AND ILANA HADAR	
Lysergic acid diethylamide (LSD 25): XXXX. Effect of pH on transport of methysergide and LSD 25 across gill membrane . . . . .	111
BY H. H. GETTNER, P. A. CARONE, AND H. A. ABRAMSON	
The effect of experimenter absence and response delay on nonreinforced imitation . . . . .	119
BY PAUL M. SMEETS AND SEBASTIAN STRIEFEL	
An investigation of the relationship between stimulus explicitness and entering behavior in facilitating student achievement . . . . .	129
BY THOMAS C. ARNOLD AND FRANCIS M. DWYER	
An exploratory study of individual differences in perceptual centering and de-centering . . . . .	133
BY ARIE COHEN AND FRANK H. FARLEY	
Creative thinking ability, school readiness, and intelligence in first grade children . . . . .	137
BY NICHOLAS C. ALIOTTI AND WILLIAM E. BLANTON	
A critique of Sarnoff and Zimbardo's psychoanalytic alternative to a social comparison theory of emotions . . . . .	145
BY PAUL L. WUEBBEN	
Fertility ratio: Its relationship to mental ability, school achievement, and race . . . . .	159
BY R. T. OSBORNE	
Changes in college student attitudes toward the Arab-Israel, India-Pakistan, and Vietnam conflicts . . . . .	165
BY WILLIAM B. ROSENBAUM AND LEONARD L. ROSENBAUM	
Change-making strategies in children and adults . . . . .	173
BY HOWARD R. POLLIO AND ROGER T. GRAY	
Disparagement by a subordinate, ingratiation, and the use of power . . . . .	181
BY EUGENE M. FODOR	

## JUDGMENT OF RELATIVE REGENCY: DEVELOPMENTAL TRENDS\*<sup>1</sup>

*University of Turku, Finland*

---

J. M. VON WRIGHT

---

### SUMMARY

The development of the ability to judge the relative recency of items within a series of successively presented pictures was studied with a forced-choice technique. Ss were 5, 8, and 12 year old children, and adults (total  $n = 239$ ). In group CA 5, accuracy of judgment was clearly above chance level only when the second (more recent) picture was inside and the first one well outside the immediate memory span. The main improvement occurred between CA 5 and CA 8; after CA 8 only judgments of pairs in which the second picture immediately preceded the test continued to improve.

In a second experiment with 5, 8, and 12 year old Ss (total  $n = 160$ ), a repeated (twice shown) picture tended to be judged more recent than a once shown one when both pictures were outside the immediate memory span, when the repeated occurrences were massed (but not when they were spaced), and when the Ss were 8 years or older. When making erroneous judgments of recency, the Ss in all age groups tended to judge that picture of a pair more recent which was also judged to be more interesting. The results are in agreement with the hypothesis that trace strength is a variable affecting judgments of relative recency. Some possible reasons for the low level of performance of the youngest Ss were briefly considered.

---

### A. INTRODUCTION

Judgments of the recency of past experiences are frequent in everyday life, but the process underlying these judgments is not very well understood. It involves memory, and it frequently involves a fair amount of reasoning (e.g., "It must have been before Christmas, because . . .") and decision processes. In addition, absolute judgments of recency (JOR) usually involve time esti-

---

\* Received in the Editorial Office on January 29, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This study was supported by the Research Council for the Social Sciences and the Jenny and Antti Wihuri Foundation. The author is grateful to Mrs. Pirkko Wük for assistance in the data collection.

mation. Judgments of the relative recency of two or more experiences may be simpler, since they presuppose only a discrimination of recency but no articulated time-related scale in terms of which the judgment is made. Thus there is abundant anecdotal evidence of children's ability to make correct relative JORs prior to the age of five (e.g., 10, 11), an age at which reasonably accurate absolute JORs are an exception [see Fraisse (3)]. However, apparently there are no systematic studies of the development of children's ability to make JORs.

Most experimental studies of JOR have been concerned with judgments of items within a continuously presented series of words or pictures, the maximal time span being 20 minutes or less. This reflects the fact that the study of JOR has been linked with the study of short-term memory processes (4). The research has mainly focussed on the nature of the cues involved in recency discrimination. Konorski (5) has suggested that this discrimination is based on a comparison of the strengths of the memory traces involved. This suggestion receives some support from studies showing that repeated events outside the immediate memory span tend to be judged more recent than events occurring only once, provided that the repeated occurrences are massed rather than spaced (7, 8). It has also been incorporated in the "two-process memory-strength theory" for judgment of recency advanced by Hinrichs (4). However, trace strength theories have a rather limited applicability, as pointed out by Tulving and Madigan (12): in a variety of situations the empirically found JORs are largely dependent on "time-tags" unrelated to trace strength.

The present study is primarily explorative: it was thought to be of interest to obtain some information about the ontogenetic development of the ability to make JORs. It is well known that performance based on simple recognition memory develops rapidly. For instance, Brown and Scott (1) found the ability of 3-5 year old children to recognize old pictures to approximate that of adults. However, when a memory task requires the ordering of the remembered events according to some temporal or spatial rule, a gradual improvement of performance with increasing age of the Ss is often evident long after the age of seven, the rate of the developmental change depending on the nature and the complexity of the rule involved (2, 6, 9). The rule involved in the ordering of two events in terms of their relative recency is comparatively simple. Some pilot experiments indicated that 5 year old children could easily be taught to make correct judgments of relative recency when three successive pictures were presented, whereas difficulties were often encountered and special training procedures required with 4 year old Ss. The youngest age group in the present study is CA 5.

A simple method for studying recency discrimination is to employ sets of three successive stimulus events: events A and B, and a forced-choice test T, in which *S* is asked to compare A and B in terms of their recency of past occurrence. In this type of task, two time intervals can be varied independently:  $A \rightarrow B$  and  $B \rightarrow T$ . In what appears to be the first experiment employing this method, Yntema and Trask (13), using adult *Ss*, found the accuracy of JOR to vary directly with the interval  $A \rightarrow B$ , and inversely with the interval  $B \rightarrow T$ .

## B. EXPERIMENT 1

### 1. Method

Each *S* was shown successively 116 cards for 3 seconds each with an interval of 2 seconds between the cards. About two-thirds were inspection cards: a color photograph (6 by 9 cm) of a common object, pasted in the middle of a 15 by 20 cm piece of white cardboard. The other cards were test cards with two such pictures pasted side by side. *S* was to judge which of the two pictures he had seen more recently and to point at that picture, his response being recorded by *E*. A picture would appear at most twice—on an inspection card and then on a test card—and *S* was told of this fact.

Essentially the series consisted of three-card subsets: inspection cards A and B, and test card T containing pictures A and B. The *distance* between any two pictures (cards) in the whole series is defined as the number of intervening pictures (a test card being counted as two pictures). In the experiment two distances,  $A \rightarrow B$  and  $B \rightarrow T$ , were independently varied, both being given four values: 0, 3, 7, and 13. There were thus a total of  $4 \times 4 = 16$  different types of subsets. An illustration: subset  $\{A \rightarrow B = 3, B \rightarrow T = 7\}$  refers to the sequence "A, x, x, x, B, x, x, x, x, x, x, T," where "x" denotes a picture not in the set. There were two instances of each subset, one occurring in the first and one in the second half of the series. In addition there were 20 filler cards (11 inspection and six test cards), seven of which were used as initial cards in the series.

Two series, identical in structure but with different pictures, were used, Series I being presented to half of the *Ss*, and Series II to the other half. The experiment was performed individually with each *S*. In connection with the instructions a brief preliminary task was presented in order to ensure that *S* understood the task, the criterion for understanding being that *S* twice in succession made the correct choice in a miniature JOR task involving four presentations of objects (task type: "A, x, B, T").



## 2. Subjects

Subjects were 5 year old children, age range 61-72 months, from kindergartens ( $n = 48$ ); 8 year old children from the second grade in elementary schools ( $n = 49$ ); 12 year old children partly from the sixth grade in citizen's schools ( $n = 45$ ), partly from the second grade in secondary schools ( $n = 57$ ); and adults, age range 19-25 years, partly university students ( $n = 20$ ), partly warehouse employees ( $n = 20$ ). Each group contained an approximately equal number of male and female Ss.

## 3. Results

The pooled results for each age group are presented in Table 1 and Figure 1. Since a forced-choice test was used, the results are expressed as proportions of correct choices. It can be seen that in the youngest age group, CA 5, the proportions fluctuated around the .50 level in all cases but two. Performance was significantly ( $p < .01$ , as evaluated by a  $t$  test for independent proportions) better than expected on the hypothesis of random guessing when  $B \rightarrow T = 0$  and  $A \rightarrow B = 7$  or 13; when, in other words, the latter picture was well within the immediate memory span and the former picture (A) well outside it.

TABLE 1  
PROPORTION OF CORRECT CHOICES IN EACH CONDITION IN EACH AGE GROUP

Age group	A $\rightarrow$ B	B $\rightarrow$ T			
		0	3	7	13
CA 5 ( $n = 96$ )	0	.55	.54	.45	.53
	3	.57	.46	.58	.53
	7	.72**	.54	.46	.54
	13	.73**	.53	.60*	.52
CA 8 ( $n = 98$ )	0	.86**	.57	.53	.53
	3	.80**	.66**	.63*	.64**
	7	.83**	.67**	.62*	.61*
	13	.81**	.76**	.69**	.71**
CA 12 ( $n = 204$ )	0	.91**	.58*	.57*	.57*
	3	.93**	.78**	.58*	.61**
	7	.93**	.72**	.69**	.68**
	13	.92**	.79**	.75**	.76**
Adults ( $n = 80$ )	0	.92**	.55	.55	.57
	3	.96**	.74**	.56	.57
	7	.96**	.72**	.65**	.62*
	13	.98**	.72**	.70**	.72**

Note: A  $\rightarrow$  B = distance interval between cards A and B; B  $\rightarrow$  T = distance interval between card B and the test card.  $n$  = number of observations from which the proportions were calculated.

\* Tabled proportion  $> .50$  at the  $p < .05$  level.

\*\* Tabled proportion  $> .50$  at the  $p < .01$  level.

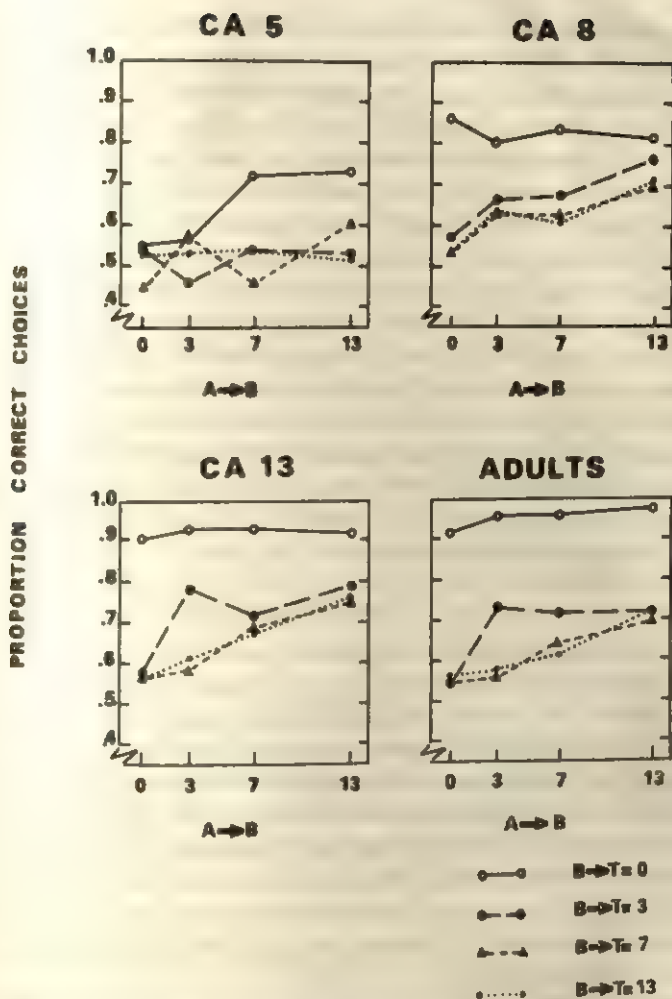


FIGURE 1

PROPORTION OF CORRECT CHOICES IN EACH CONDITION IN EACH AGE GROUP

In the other age groups the performance level was substantially higher. The main improvement occurred between CA 5 and CA 8. In fact, consistent or clear-cut evidence for improvement of performance *after* CA 8 was found only in the cases in which  $B \rightarrow T = 0$ . This result is of some interest, since it seems to suggest that—in the present type of task—Ss become increasingly more able with increasing age to use information in short-term memory for rapid judgments of recency, whereas no corresponding improvement occurs when

performance is dependent on information that has been "in store" for some time (distance  $B \rightarrow T = 3$  corresponded to a time interval of 15 seconds).

When the effects of the variation of the time interval  $A \rightarrow B$  and  $B \rightarrow T$  are compared, the pattern of results is similar for all age groups except CA 5. When  $B \rightarrow T$  increased, performance level decreased rapidly at first, but this decrease leveled off gradually: the results for  $B \rightarrow T = 7$  and for  $B \rightarrow T = 13$  are virtually identical (see Figure 1). This agrees with commonsense expectations: once the latter event is well outside the span of immediate memory, it makes little difference if it occurred 30 seconds or 1 minute ago. An increase in the temporal distance between the two events to be compared ( $A \rightarrow B$ ), on the other hand, made for a small but fairly consistent improvement over the whole range of distances used in the present study (at least when  $B \rightarrow T > 3$ ).

In each age group there was a small but fairly consistent sex difference in performance level in favor of the female Ss. However, the average difference reached statistical significance ( $p < .01$ ) only in the youngest age group, in which the girls achieved a higher proportion of correct responses than the boys in 13 of the 16 stimulus subsets.

### C. EXPERIMENT 2

In a second experiment an attempt was made to study some factors determining judgments of relative recency. As mentioned in the introduction, "trace strength" has usually been considered a construct of special interest. One might even argue that trace strength should be of greater importance for JOR in younger children than in older ones on the grounds that the latter are more likely to use systematic approaches to the task: e.g., the use mnemonic devices serving to attach "time-tags" to the stimulus events. Trace strength was varied by presenting both pictures in a pair (A and B) either once or twice, the repeated occurrences being either massed or spaced [for a thorough discussion of the expected differences between massed and spaced presentations, see Peterson *et al.* (8)].

In addition, an attempt was made to ascertain how interesting the Ss considered the pictures in each test pair to be. One might expect that the more interesting an event is experienced to be, the more time is spent in studying it and/or the more attention is focussed on it; hence, its trace strength should be greater than that of a less interesting event.

#### 1. Method

The experiment was identical to Experiment 1 except that the series of items shown to the Ss had a different structure. Basically there were four

types of item subsets, obtained by combining in all possible ways two values of  $A \rightarrow B$  and  $B \rightarrow T$ :  $\{A \rightarrow B = 3, B \rightarrow T = 3\}$ ,  $\{A \rightarrow B = 3, B \rightarrow T = 13\}$ ,  $\{A \rightarrow B = 13, B \rightarrow T = 3\}$ ,  $\{A \rightarrow B = 13, B \rightarrow T = 13\}$ . Five variants of each subset were constructed as follows: (a) A and B both occurred once. (b) A occurred twice massed (distance between occurrences = 0) and B once. (c) A occurred twice spaced (distance between occurrences = 15) and B once. (d) A occurred once and B twice massed. (e) A occurred once and B twice spaced (a distance of 15 between the occurrences implying that B in fact occurred once before A and once after it).

Within the whole series, S thus made  $4 \times 5 = 20$  judgments of relative recency which were included in the results. In addition the series contained 16 filler cards (11 inspection and five test cards), six of which were used as initial cards in the series. Two series, identical in structure but with different pictures, were used, Series I being presented to half of the Ss and Series II to the other half. The experiment was performed individually with each S.

After the JOR experiment, S was shown each pair of pictures which in the series had appeared on the same test card, one pair at a time in a random order. He was asked to point at that picture which he thought was more interesting to look at or which he liked better. The responses were recorded by E.

## 2. Subjects

Subjects were 5 year old children (age range 61-72 months) from kindergartens ( $n = 40$ ); 8 year old children from the second grade in elementary schools ( $n = 40$ ); and 12 year old children partly from the sixth grade in citizen's schools ( $n = 40$ ), partly from the second grade in secondary schools ( $n = 40$ ). Each group contained an approximately equal number of boys and girls.

## 3. Results

The effect of repeated occurrences of items was evaluated by comparing, within each item subset, the proportion of choice of the repeated item (in variants *b*, *c*, *d*, and *e*) with the proportion of choice of the corresponding item in variant *a*, in which both A and B occurred only once. Since the *a*-variants corresponded to some of the conditions in Experiment 1, the corresponding data from the two experiments were pooled in order to obtain "standards" based on a larger number of observations.

This treatment of the data yields 16 comparisons (differences) for each age group. In the youngest age group, CA 5, the effects of repeated occur-

rences were negligible, none of the 16 differences reaching the  $p = .05$  level of significance. Similar negative results were obtained in the two older age groups when the repetitions were spaced (variants *c* and *e*). Massed repetitions, however, appeared to have some effect. A massed repetition of the *earlier* item (variant *b*) led to a significant ( $p < .01$ ) increase in the proportion of wrong responses in groups CA 8 and CA 12 when  $A \rightarrow B$  was small (distance 3) and  $B \rightarrow T$  was large. Correspondingly, a massed repetition of the *latter* item (variant *d*) led to a significant ( $p < .05$ ) increase in the proportion of correct responses when both  $A \rightarrow B$  and  $B \rightarrow T$  were large (distance 13).

As far as the results go, they suggest that there was a tendency for the repeated item to be judged more recent when the repetitions were massed, when both pictures were well outside the immediate memory span ( $B \rightarrow T = 13$ ), and when the Ss were 8 years or older. The results obtained with the older Ss are thus similar to those obtained by Peterson *et al.* (8) with adult Ss and a very different technique, though Peterson's results were definitely more clear-cut than the present ones. The hypothesis that the effects of item repetition are stronger with younger Ss than with older ones, on the other hand, received no support; the data indeed suggest the opposite conclusion.

Each pair of pictures occurring together on a test card was (after the JOR experiment) rated in terms of which of the pictures was the more interesting one to look at. A slight relationship between these "preference" judgments and the JORs is evident in the overall data: in all age groups the proportion of preferred pictures is somewhat higher in the set of pictures judged more recent than in the set of pictures judged less recent. A further analysis shows that this relationship is absent when the *correct* recency judgments alone are considered; it is, however, significantly present in the set of *wrong* recency judgments. When making erroneous JORs, the Ss in group CA 5 judged the more interesting picture to have occurred more recently in 59.1 percent of the cases (218 out of 369 errors). The corresponding percentages are 57 percent (131 out of 230 errors) in group CA 8, and 62.7 percent (244 out of 389 errors) in group CA 12. This would seem to suggest that the Ss, when uncertain, tended to judge the more interesting picture to be the more recent one. One interpretation of this result is in terms of trace strength: when confronted with a long sequence of pictures, S may tend to spend more time watching or thinking about a picture which he considers comparatively interesting than one that arouses less interest. The more interesting picture stands a better chance to be efficiently memorized.



#### 4. Discussion

The results lend some support to the trace strength hypothesis for judgment of recency as follows: (a) In groups CA 8 and CA 12 repeated events outside the immediate memory span tended to be judged more recent when the repeated occurrences were massed but not when they were spaced, and (b) the Ss in each age group tended to judge the more interesting picture of a pair more recent than the less interesting one when making an erroneous JOR. However, even when statistically significant, these tendencies were fairly weak. This suggests that only a modest part of the Ss' response variance in the task can be accounted for in terms of differential trace strength.

The main bulk of data obtained with 5 year old children is consonant with the hypothesis that the Ss responded by "random guessing" unless the latter picture (B) immediately preceded the test and the distance between pictures A and B was large. As stated above, the Ss appeared to understand the task requirement, as evidenced by the fact that they made correct JORs in "miniature" tasks; as far as *E* could judge, their motivation and interest in the task also seemed to be satisfactory. Hence the apparent randomness of their performance may be worth a closer analysis: what is the nature of the limitations to the ability of 5 year old children to judge relative recency in tasks of the present type? The data do not contain definite clues to the solution of the problem. However, it seems likely that the length of the series is a task variable of some importance. Studies of the development of attentional processes indicate that younger children are less selective in their information intake in learning tasks than older ones. Thus they are less likely to focus their attention continuously on those features of the situation that are relevant to the performance of the *E*-defined task. The longer the series, the more detrimental to performance are shifts of attention from task-relevant to task-irrelevant features. Older children are also more likely than younger ones to develop "strategies" that serve to reduce intratask interference. As an example of such a strategy may be mentioned the attempts at "intentional forgetting" of pictures that had appeared on a test card, occasionally mentioned by adult Ss who were interviewed about how they tried to cope with the task. Hence the age differences in accuracy of JOR may be expected to decrease when the length of the series of items presented decreases, and possibly also when intratask interference is reduced (e.g., by decreasing the similarity of the items in the series).

## REFERENCES

1. BROWN, A. L., & SCOTT, M. S. Recognition memory for pictures in preschool children. *J. Exper. Child Psychol.*, 1971, **11**, 401-412.
2. BRUNER, J. S., OLVER, R. R., GREENFIELD, P. M., *et al.* *Studies in Cognitive Growth*. New York: Wiley, 1966.
3. FRAISSE, P. *The Psychology of Time*. New York: Harper & Row, 1963.
4. HINRICHES, J. V. A two-process memory-strength theory for judgment of recency. *Psychol. Rev.*, 1970, **77**, 223-233.
5. KONORSKI, J. The physiological approach to the problem of recent memory. In J. F. Delafresnaye (Ed.), *Brain Mechanisms and Learning*. London: Oxford Univ. Press, 1961. Pp. 115-132.
6. NELSON, K. E. Memory development in children: Evidence from nonverbal tasks. *Psychon. Sci.*, 1971, **25**, 346-348.
7. PETERSON, L. R. Search and judgment in memory. In B. Kleinmuntz (Ed.), *Concepts and the Structure of Memory*. New York: Wiley, 1967. Pp. 153-180.
8. PETERSON, L. R., JOHNSON, S. T., & COATNEY, R. The effect of repeated occurrences on judgments of recency. *J. Verb. Learn. & Verb. Behav.*, 1969, **8**, 591-596.
9. PIAGET, J., & INHELDER, B. *Mémoire et Intelligence*. Paris: Presses Univ. France, 1968.
10. STERN, C., & STERN, W. *Erinnerung, Aussage und Lüge in den Frühen Kindheit* (4th ed.). Leipzig: Barth, 1931.
11. STERN, W. *Psychologie der Frühen Kindheit* (5th ed.). Leipzig: Quelle & Meyer, 1928.
12. TULVING, E., & MADIGAN, S. A. Memory and verbal learning. *Ann. Rev. Psychol.*, 1970, **21**, 437-484.
13. YNTEMA, D. B., & TRASK, F. P. Recall as a search process. *J. Verb. Learn. & Verb. Behav.*, 1963, **2**, 65-74.

*Department of Psychology*

*University of Turku*

*Yliopistonkatu 23*

*20500 Turku*

*Finland*

## MANIPULATION OF SITUATIONAL FAVORABILITY IN TESTS OF THE CONTINGENCY MODEL\*

*State University of New York at Buffalo*

---

L. H. GEYER AND J. W. JULIAN

---

### SUMMARY

Two experimental studies were performed which directly manipulated the situational favorability for the group leader by controlled interference with his ability to communicate with his group. In the first experiment subjects were industrial supervisors; in the second, university students. Evidence indicated that the experimental manipulation was seen as influencing situational favorability. Results obtained provide a measure of support for the contingency model of leadership effectiveness. Experiment 2 utilized an anagram game in which it was possible to parcel out variance resulting from the task-relevant ability of group members.

---

### A. INTRODUCTION

Fiedler's (4) contingency theory of leader effectiveness focuses on the dimension of "situational favorableness" as the variable that moderates the relationship between group performance effectiveness and leader LPC. LPC stands for the leader's affective rating of his "least preferred co-worker"; it presumably indexes the degree to which he values the interpersonal or task dimensions of group member performance (4). The contingency model predicts that situational favorableness moderates the relationship between leader LPC and effectiveness in such a way that the relationship is negative when favorableness is either very high or very poor, and it is positive for conditions that are merely moderately favorable.

In his 1967 monograph, Fiedler defined situational favorability as "... to what extent the situation will make it easy or difficult for the leader to influence his group members" (4, p. 142). Previous research (2, 3), however, operationally defined situational favorability in terms of an ordered hierarchy of specified variables: leader-member relations, task structure, and leader position power. Thus, while the definition was intuitively appealing, it has

---

\* Received in the Editorial Office on February 6, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

not been clear what combinatorial rule should be used to map these variables to the favorability dimension, nor has previous research tested directly the fit between the definition of favorability and its presumed elements.

The present report presents two investigations representing a new tack by manipulating directly the favorability of the situation by altering the leader's ease of interaction with his group in a way that can be specified quantitatively. The proposed methodology, therefore, would at last offer the prospect of exploring the hypothesized LPC-effectiveness relationship at selected points along the favorability dimension in a way that could be replicated across a variety of industrial or laboratory task settings.

The manipulation of favorableness was based on the assumption that if the leader were restrained to silence, literally held incommunicado, during portions of the group's efforts at an interactive task, then such a restriction would disadvantage his ability to exert influence. The degree to which his ease of influence was impaired was assumed to increase monotonically with the proportion of group interaction time that he was held incommunicado. Study 1 explores the hypothesized moderating effects of the favorableness dimension in an industrial setting. Study 2 explores the same hypothesis in a laboratory setting, using a different task and different population.

## B. EXPERIMENT 1: INDUSTRIAL SETTING

### 1. Method

*a. Subjects and design.* One hundred four volunteers from the supervisory staff of a suburban industrial concern participated as members of 26, four-person groups. The first investigator was at the time a supervisory employee of this company. In recruiting Ss, he informed them that the study, in addition to potential relevance to management theory, was directly a portion of his graduate study program. The avowed purpose of the experiment was "to study communications effects on group problem solving." Each group met in its occupational milieu for a single 90-minute discussion session. The members were randomly assembled with the one constraint that supervisors were not grouped with their actual subordinates. Each group was randomly assigned to one of three "communication interruption" schedules. These schedules comprised the manipulation of the favorability of the situation for leader influence. All groups completed two discussion tasks in counterbalanced order. The design thus permitted an examination of the relationship of leader LPC to group effectiveness under each of three favorability conditions and for each of the two tasks.



*b. Procedure.* The procedure for each session was quite simple. Initially, members were given written introductory material which explained the nature of the problem solving task, the manipulation of communications, and the overall time constraints on performance.

All Ss completed LPC ratings during the first part of the session, after which a "chairman" was selected by rolling dice. The LPC instrument used [see Fiedler (4), Appendix A] consisted of 16 bipolar pairs of adjectives: e.g., pleasant/unpleasant. S was instructed to think of "the person with whom you can work least well," and to rate that person from 1 to 8 on each of the 16 items. The LPC score was the arithmetic sum of the 16 ratings. Responsibilities and prerogatives of the chairman were then explained, and he was given the first problem statement to read. When he indicated readiness, the first 30-minute discussion period began. Following the completion of problem one, this procedure was repeated for the discussion of the second problem. "Reports" were written by the chairman during the allotted 30-minute period. The experimenter always provided a five-minute warning. Finally, all Ss filled out ratings of the climate or atmosphere of their groups on a set of evaluative semantic differential scales and a questionnaire designed to assess the effectiveness of the manipulation of leader communication.

*c. Favorability of the leader's position.* The aims of this investigation imply the manipulation of the favorability of the leader's situation from some known background condition of favorability. For Experiment 1, the background setting that anchored the manipulation was one in which group-centered decision making (6) was a positively valued corporate way of life. In this environment, hierarchically heterogeneous problem solving groups were well practiced, familiar, and structured by a conscious tradition. Hence, the leadership position of the temporary chairman of such groups was well established. The group members were, collectively, about three-quarters of the total supervisory force of the organization. Hence they were in frequent and normally friendly interaction. These observations suggest that the initial condition, from which manipulation of situational favorability would be attempted, was a favorable one.

To manipulate the favorability of the situation for the leader to exert influence, three conditions of communication restraint were created. First, each problem discussion period was thought of as 60 one-half-minute intervals. Intervals of enforced silence for any individual were signaled by the lighting of a small signal light placed directly in front of him. For a favorable, low interruption, condition the leader was constrained by these signals not to talk during 5 percent (i.e., three) of the 60 intervals. For a less favorable,



medium interruption, condition, he was constrained from talking 15 percent of the time, and for an unfavorable, high interruption, condition, he was constrained for a total of 30 percent of the group discussion time. For non-leaders in all conditions total interruption time was 10 percent of the 60 intervals. The specific schedules of enforced silence were preselected on a random basis, except that no more than one member could be held incommunicado at a time.

*d. Discussion tasks.* Two tasks were used, each requiring the group to discuss a problem drawn from their industrial organization and to report their recommendations in writing. Group members were informed that their report would be the sole basis on which performance would be judged. Contrasting problems were chosen which varied the "task structure" dimension of Fiedler's contingency model (4, pp. 25-28). The "structured" problem posed the challenge of improving supervisory training within the plant, while the "unstructured" problem dealt with the long range issue of establishing policy regarding corporate citizenship, specifically the participation of corporate employees in community affairs. Four judges (three corporate officers and the corporate personnel consultant) rank ordered the group reports for each problem separately on the basis of criteria presented in the problem statements given to the groups. Average interjudge correlation was .56 for the structured problem and .31 for the unstructured. Individual judge rankings were combined to arrive at an overall average score which was used for analysis. An apparent "position effect" favoring the second discussion over the first emerged, which was of no theoretical interest. Group performance scores were adjusted for this position effect before analysis.

## 2. Results

There were no systematic differences in the scores on the structured *vs.* unstructured tasks. The rank order correlation between leader LPC and group performance, averaged over both tasks, for the low interruption condition was  $-.53$  ( $N = 8$ ); for the medium interruption condition  $\rho = -.20$  ( $N = 9$ ); for the high interruption condition  $\rho = +.25$  ( $N = 9$ ). It is apparent that these effects of the communication interruptions on the LPC-group effectiveness relationship were monotonically increasing, rather than the anticipated curvilinear function. The difference between coefficients at extreme conditions was significant ( $p < .05$ , one tailed test). The three communication interruption conditions were intended to vary the favorability of the leader's position for influencing the group. The effectiveness of these interruptions was assessed by the leader's reaction on postsession ratings to the

question: "To what degree did you find this interruption or silencing frustrating?" The mean ratings of frustration were 8.5 for the low interruption condition, 11.0 for the medium interruption, and 11.8 for the high interruption condition. The difference between low and high interruption conditions was significant,  $t(15) = 2.36$ ;  $p < .05$ . This perception of frustration did not generalize to perception of group atmosphere; leaders in the high interruption condition rated the group atmosphere slightly higher than did low interruption condition leaders (the difference did not approach significance,  $t = .57$ ;  $p > .50$ ).

### C. EXPERIMENT 2: LABORATORY SETTING

One obvious interpretation of the results of Experiment 1 was that the particular experimental interruption schedule sampled too favorable a range along the scale, and hence the contingency model's recursion to negative correlation coefficients was not manifest. The magnitude of the communication restriction in each case was, in fact, arbitrary, since Experiment 1 was the first effort to manipulate directly the favorability dimension in this way. Experiment 2 was intended to extend the range along the leader favorability dimension by using the communication interruption manipulations under test conditions requiring greater leader coordination and decision making.

#### 1. Method

*a. The team anagram game.* This game evolved over several iterations; only the final version is described. Teams are seated, one player to a side, about a square game board. Below the top surface of the game board along each side are receptacles holding two sets of 10 plastic capital letters. At the start of a game, each player removes the 10 letters from his upper receptacle and places them in front of him on the top surface of the game board, so that all 40 letters are in view of all four players. The first of five timed "rounds" of two minutes each begins immediately.

At the end of each round, one letter from each player's set of 10 is played into a common kitty in the center of the game board, and then one letter from that kitty is returned to each player. The two-minute time period is for team discussion of the most advantageous set of exchanges. At the end of the fifth round each player uses as many as he is able of the 10 letters then in front of him in forming words of at least four letters in length. Letters used in words count toward team score. Letters are color coded, designating particular letters that count extra if used; J, Q, X, and Z are each worth three ordinary letters; B, V, and G are worth two. Team score is the sum of all

letters played into words times their scoring value. The scoring system was designed to emphasize team rather than individual performance. The printed game instructions pointed out that sometimes "sacrifices" by one player to help another could increase team score. The second receptacle of 10 letters before each player constitutes a "second half" of the game identical to the first, except for the different letters.

At the end of each two-minute round all players must cease talking or otherwise signaling. It is the responsibility of the team captain to select which letter from each player's set of 10 will be placed in the kitty and which letter from the kitty will be returned to each player. Regardless of discussions during the round, actual exchange at round's end is the captain's prerogative.

*b. Relationship of member ability to team performance.* Performance data from 21 four-person groups were used to develop a procedure for parceling out that proportion of team performance that was a function of member ability. First, a five-minute timed test was administered which required Ss to form as many crossword puzzle type squares of three letters each as they could. Total number of letters used, scored as in the anagram game, was the individual ability measure. Multiple regression analyses were performed which considered as potential covariants each player's individual ability score singly and in combination. In general, these analyses indicated that the individual ability measure could account for nearly 25 percent of the variance in team scores. The ability of the player to the captain's left, times the captain, and the ability of the player to the captain's right, times the captain, were dominant as individual correlates. This combination resulted in a correlation of .43 with team performance. This was considered to be a potentially useful apportionment of the team score variance not related to the stylistic features of the leader's performance.

*c. Subjects and procedure.* One hundred twenty-four undergraduate volunteers participated in 31 four-person groups, thereby fulfilling a course requirement. At the start of each session Ss completed the individual anagram task and LPC instrument and read the description of the team anagram game. Then each team played a practice set of four one-minute rounds, after reading the rules and discussing questions with E. Finally, each team selected its captain by confidential written ballot. The practice rounds were included in the procedure to ensure that all Ss understood the game and had some informed basis for voting for the captain.

Fifteen groups were assigned randomly to a minimal interruption condition (5 percent), and 16 groups to a much interruption condition (30 percent). To create the interruption schedules, each two-minute round was thought of

as four half-minute intervals, making 20 time periods per half, or 40 periods per game. In the minimal interruption condition each player, including the captain, received two periods of required silence. In the much interruption condition the captain received 12 periods of required silence. Other players in both conditions were restrained for two periods. The interrupted periods were assigned at random, except that no more than one player could be held silent at a time. The silent periods were signaled by individual lights behind translucent panels below the top surface of the game boards on each side. Occasionally *E* had to call attention to a light during active play.

*d. Favorability of the leader's position.* As in Experiment 1, an assessment of a background level of favorability is implied. In this case group members had no prior shared experience of participation in any more particularized subculture than that of a large university. Within this context, the leader's position power seemed fairly described as weak, and his relations with other group members as approaching neutral. Although he was elected to office, it was under rather forced and transient conditions. The degree of task structure was also intermediate; goal clarity was high and decision verifiability reasonably high, but goal path multiplicity was quite large, and/or solution specificity was thereby low. In addition, the context of the laboratory experience appeared to induce at least moderate levels of "evaluation apprehension" among college student subjects (7). These conditions seem best described as intermediate in situational favorability.

## 2. Results

It was obvious from observation that the game achieved the objective of being enjoyable and challenging. Ss remained after many sessions to discuss strategy with one another, and in some cases *E* remained to replay some portion of the game using a different plan. Some Ss also returned at a later date to find out how their team scores compared to other groups. There were frequent commentaries on the commercial potentials of the game.

The covariation of team performance with the combination of individual ability scores was quite similar in the present sample to that obtained in the relationship of member ability to team performance study described earlier,  $r = +.48$  in the present case ( $p < .01$ ) and  $+.43$  in the pilot. Corrected performance scores consisting of the residuals from this regression analysis were used for the analysis of leader LPC to group performance relationships. The averages of these corrected scores did not appear to indicate a differential induction due to communication conditions. The difference in performance between condition means did not approach significance.



The basic aim of the second experiment was to demonstrate the feasibility of directly manipulating the favorableness of the influence situation by holding the captain incommunicado during the team's discussion of game strategy. Our hypothesis was that a minimal interruption schedule would present a condition of intermediate favorability for leader influence, while the 30 percent interruption condition would be a distinctly unfavorable setting. Thus, in terms of the contingency model, we had expected a positive relationship between LPC and team performance under the 5 percent interruption schedule, and a negative relationship under the 30 percent schedule of interruptions. The obtained rank order correlation of leader LPC with corrected group performance scores was  $+.03$  in the 5 percent interruption condition and  $-.48$  in the 30 percent interruption condition ( $p < .05$ , single tail test). The difference in these coefficients approached but did not reach significance ( $.05 < p < .10$ , single tail test).

#### D. DISCUSSION

Graen, Alveres, Orris, and Martella (5) have pointed out two critical areas for the research supporting the contingency model of leadership effectiveness. First, the usual standards of statistical reliability have been relaxed in interpreting many of these studies. Second, the "favorability" or "ease of influence" dimension is poorly defined and subject to arbitrary manipulation. The latter problem is particularly crucial because the hypothesized curvilinear relationship of the contingency model can lead to opposite expectations for the LPC to effectiveness relation depending upon the assumed positions along the favorability (ease of influence) dimension.

These issues are relevant to the studies reported. The basis for the assessment of initial condition of favorability has been set forth in the preceding text, so that the reader can formulate his own judgment as to whether or not the assessment seems valid. Beyond that the investigators can report the judgment from observation that the communication interruption manipulation was indeed making a difference in the leader's situation. Further, differences in obtained correlations of leader LPC to group performance would not appear to be attributable to a sampling artifact of differences in LPC between experimental groups. The median leader LPC scores for the extreme conditions in Experiment 1 were 59.9 and 64.9; in Experiment 2 they were 54.6 and 59.0. In neither case did the difference approach significance.

On a methodological note, some commentary on the utility of the team anagram game seems appropriate. It was very well received by Ss as an experimental task. The evidence from both the pilot study and the second



experiment indicates that the contribution of individual ability to team score is predictable. This indication is strengthened by data from an unpublished study by Enggist, Salvendy, and Geyer (1). These investigators used a three person team variation of the anagram game and the same measure of individual ability reported here. There was no captain; exchanges were consensual. A covariate function formed by summing the coupled (product) ability scores of each pairing of the triad correlated with team score ( $r = .45$ ;  $N = 18$ ;  $p < .05$  single tail test). The portion of the total sum square variance predicted by coupled ability scores was quite similar in all three studies: 18.5, 23, and 20 percent, respectively.

## REFERENCES

1. ENGGIST, R., SALVENDY, G., & GEYER, L. H. Influence of forced *versus* freely chosen pace on task performance in a social situation. Unpublished manuscript. Department of Industrial Engineering, State University of New York at Buffalo.
2. FIEDLER, F. E. *Leader Attitudes and Group Effectiveness*. Urbana, Ill.: Univ. Illinois Press, 1958.
3. ———. A contingency model of leadership effectiveness. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 1). New York: Academic Press, 1964, Pp. 149-190.
4. ———. *A Theory of Leadership Effectiveness*. New York: McGraw-Hill, 1967.
5. GRAEN, G. B., ALVERES, K. M., ORRIS, J. B., & MARTELLA, J. A. Contingency model of leadership effectiveness: Antecedent and evidential results. *Psychol. Bull.*, 1970, **74**, 285-296.
6. LIKERT, R. *The Human Organization: Its Management and Value*. New York: McGraw-Hill, 1967.
7. WEBER, S. J. Subject effects in laboratory research: An examination of subject roles, Demand characteristics, and valid inference. *Psychol. Bull.*, 1972, **77**, 273-295.

Department of Industrial Engineering  
State University of New York at Buffalo  
Parker Engineering  
Buffalo, New York 14214



## THE EFFECTS OF A ZERO INTERVAL ON SEMANTIC DIFFERENTIAL ROTATED FACTOR LOADINGS\*

*Department of Psychology and Philosophy, San Antonio College*

---

JOHN H. FORTHMAN<sup>1</sup>

---

### SUMMARY

A partial replication and extension of the semantic differential was used to test the effects of modified middle interval instructions on rotated factor loadings derived from Ss' responses. The partial replication is in good agreement with the findings published by Osgood *et al.* (5). However, the results from Ss responding to the modified middle interval instructions differ markedly from those in published and control conditions, yielding an interpretable fourth factor suggesting the achievement of more subtle nuance of meaning.

---

### A. INTRODUCTION

The semantic differential was developed to its present form by Osgood and Suci (4) and Osgood, Suci, and Tannenbaum (5). Both its reproducibility and the facility with which it can be designed to measure different aspects of meaning make it one of the most versatile and sensitive instruments of measure available to behavioral science.

On occasions, investigators using the semantic differential have reported findings that fail to produce—or reproduce—factor loadings of expected magnitude. Triandis and Osgood (10) found differences on some scales, as well as on some concepts, between monolingual Greek and American S populations. They explained the variations in a logical manner, focusing on the differences between Greek and American culture. Suci (9), studying Navajo, Hopi, Zuni, and Spanish speaking natives of the American Southwest, found that the first three factors accounted for only 39.9% to 64.1% of the common variance for the Indian Ss. Osgood *et al.* (5) indicate that the first three factors should account for 75% or more of the common variance.

Plakovitz (6), in measuring the attitudinal differences among Jews, obtained frequency distributions in the form of stylized "W"s. Dechter (1),

---

\* Received in the Editorial Office on February 9, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> I wish to thank Professor Loren E. Davis of Texas A & I University for his aid and encouragement in the performance of this investigation.

studying role concepts in the American Southwest, appended, but did not analyze, similar frequency distributions. The frequency distributions from both of these studies indicate that the problem may be interpreted in terms of Ss' systematic and disproportionate selection of the middle scale interval.

Sherif and Sherif (7, p. 375), in a discussion of the semantic differential, comment: "Further, the interpretation of the zero [middle interval] ratings is not clear." Oetting (3) raised the same question. However, his investigation did not provide a comparative analysis with published findings, did not report factor loadings, and dealt primarily with differential responses when requirements for forced choice by Ss were removed.

The present investigation is intended to explore whether modified middle interval instructions will be associated with factor loadings that indicate the achievement of more subtle measurement of meaning. Osgood and Suci (4) and Osgood *et al.* (5) appear to have grouped three unrelated aspects of meaning as criteria for S selection of the middle interval. The S is instructed to use intervals one and seven if the concept is "very closely related" to the scale; two and six if the concept is "quite closely" related to the scale; and three and five if the concept is "only slightly" related to the scale. The suggested instructions for the middle interval are as follows (5, p. 83):

If you consider the concept to be *neutral* on the scale, both sides of the scale *equally associated* with the concept, or if the scale is *completely irrelevant*, unrelated to the concept, then you should place your check-mark in the middle space:

good \_\_\_\_\_:\_\_\_\_\_:☒:\_\_\_\_\_:<\_\_\_\_\_ bad

Under these instructions an investigator may not be able to differentiate between these three different values the S is instructed to place in the middle scale interval. The evidence suggests that the middle interval becomes a dumping ground—a garbage can—for any scale on any concept with which an S has difficulty, question, or reservation.

The purpose of the present study is to investigate the effects of modified instructions pertaining to criteria for S selection of the middle interval. In an attempt to obtain responses of equality in the middle interval, the suggested instructions of Osgood *et al.* (5) were modified to allow a zero interval for any scale perceived by an S or Ss as *neutral* or *irrelevant* to any concept or concepts. The rationale for the decision to create a single zero interval rather than two zero intervals was based on two considerations. In previous administrations some Ss were confused by the length and complexity of the suggested instructions; rather than increase the complexity and length with

two additional instructions, neutral and irrelevant were grouped together. The second consideration was that for mathematical reasons it would be preferable to multiply a violation of logic (more than one conceptual category in a single scale interval) by zero rather than by four.

The hypotheses, stated in the null for application of statistical tests, is that under conditions of suggested and modified middle interval instructions (a) there is no difference in obtained rotated factor loadings between Treatment I [instructions of Osgood *et al.*, (5)] and Treatment II (zero interval instructions as modified by the *E*), and (b) there is no difference in the frequency and range of responses between Treatment I and Treatment II.

## B. METHODOLOGY

### 1. Subjects

Ss were introductory psychology students at Texas A & I University, Kingsville, Texas. The variables of sex and ethnic identity (Mexican- and Anglo-American) were controlled by stratified random assignment of Ss to treatment group. Of the 99 Ss who cooperated in the investigation, one was randomly eliminated from Treatment II to re-equate the groups; final analysis was based on 49 Ss in each treatment group.

### 2. Instrument

The instruments used in this investigation presented the 20 concepts (LADY, BOULDER, SIN, FATHER, LAKE, SYMPHONY, RUSSIAN, FEATHER, ME, FIRE, BABY, FRAUD, GOD, COP, PATRIOT, TORNADO, SWORD, MOTHER, STATUE, and AMERICA) and 21 of the 50 scales (heavy-light, sour-sweet, good-bad, worthless-valuable, sharp-dull, strong-weak, clean-dirty, fast-slow, honest-dishonest, unpleasant-pleasant, passive-active, foul-fragrant, small-large, unfair-fair, tasty-distasteful, hard-soft, nice-awful, ugly-beautiful, kind-cruel, profane-sacred, and sweet-bitter) of Osgood *et al.* (5). Directionality of each of the adjective pairs was determined by modified randomization in accordance with the instructions of Osgood *et al.* The pages were collated in varied sequences to minimize order effects. After the stimuli were assembled, the differential instructions were added.

Treatment I presented the instructions of Osgood *et al.* (5) without change. In Treatment II the instructions<sup>2</sup> for the middle interval were modified to read as follows:

<sup>2</sup> Several logical alternatives exist for the division of these three categories. The crea-



If you consider the concept to be *equally associated* with the scale you should place your check-mark as follows:

good \_\_\_\_\_:\_\_\_\_\_:\_\_\_\_\_✓\_\_\_\_\_:\_\_\_\_\_:\_\_\_\_\_ bad

and the following instruction was added:

If you consider the concept to be *neutral* or *irrelevant* with the scale draw a line through the entire scale as follows:

good \_\_\_\_\_:\_\_\_\_\_:\_\_\_\_\_:\_\_\_\_\_:\_\_\_\_\_ bad

[Here the instruction shows a sample seven-interval scale with a line drawn through it.]

Thus, the only differences between Treatment I and Treatment II instruments were the instructions relevant to the selection of the middle interval.

### 3. Procedure

The instructor introduced the *E*, commented on the value to the student of participating in a psychological experiment, and asked for their cooperation. Administration was performed in a regular classroom and the psychology laboratory—set up as a classroom. Assignment of treatment group to location was counterbalanced to equate any effect of these variables. The *E* introduced the experiment as one in methodology and assured the *Ss* that their responses would not be attributed to them personally. The same *E* administered the instrument in both treatment conditions under closely standardized *E* behavior.

### 4. Data Analysis

The data for each *S* were punched into holograph cards with one card per concept per *S*. Each card contained a sex-ethnic code (1-4 in column 1), a treatment code (1-2 in column 3), an *S* identification code (01-*n* in columns 12/13), a concept code (01-20 in columns 23/24), and the value of each scale was punched into the odd-numbered columns 29-69. In Treatment I the scale values were punched in accordance with the instructions of Osgood and Suci (4). Treatment II data were punched the same as Treatment I data with the addition of a zero punched when an *S* drew a line through the scale indicating *neutral* or *irrelevant*. In no instance where an individual failed to respond to a scale or scales were any of the responses of that individual used in the analysis.

tion of two separate zero intervals and placement of equal and neutral in the middle interval with irrelevant in the zero interval are two of the reasonable alternatives. However, the availability of *Ss* as well as practical aspects of logistics precluded these parts of the investigation.

The analyses were performed on an IBM 360 Model 50 computer.<sup>3</sup> The following programs from the *System/360 Scientific Subroutine Package (360-CM-03) Version III* were used: the (varimax) factor analysis and the Kolmogorov-Smirnov two-sample test (2). The factor analysis for each treatment was performed in accordance with the instructions of Osgood and Suci (4) and Osgood *et al.* (5), and with the program instructed to accept only the first four factors.

### C. RESULTS<sup>4</sup>

The rotated factor loadings from Treatment I were compared with those published by Osgood *et al.* (5): 18 of 21 major factor loadings are in excellent agreement; of the three minor factor loadings for each of the 21 scales, 10 agree within  $\pm 0.20$  for all three factors, nine of the scales have similar agreement for two of the three, and two have similar agreement for one minor factor loading. The comparison of Treatment I results with those published by Osgood *et al.* (5) constitutes a substantial replication of the findings on which Treatment I was based (see Table 1).

<sup>3</sup> All data acquisition programs were written and executed by Mr. James W. Gill.

<sup>4</sup> The zero interval accounted for 10% or fewer of the Ss' responses to any scale across all concepts. Further, male Ss selected the zero interval approximately twice as frequently as did female Ss. Ethnic identity had no noticeable effect on this observation.

TABLE 1  
COMPARISON OF ROTATED FACTOR LOADINGS FROM PUBLISHED  
FINDINGS (5),<sup>a</sup> PARTIAL REPLICATION (TREATMENT I),  
AND MODIFIED INSTRUCTIONS (TREATMENT II)

Scales <sup>b</sup>	Factors				
	I	II	III	IV	$R^2$
heavy-light					
Osgood <i>et al.</i> (5)	-.36	.62	-.11	.06	.53
Treatment I	.378	.886	-.021	-.043	.866
Treatment II	.402	-.826	.214	-.102	.810
sweet-sour <sup>c</sup>					
Osgood <i>et al.</i>	.83	-.14	-.09	.02	.72
Treatment I	.854	.426	-.005	-.149	.877
Treatment II	.429	-.453	.719	-.111	.846
good-bad					
Osgood <i>et al.</i>	.88	.05	.09	.09	.79
Treatment I	-.961	-.179	.079	.043	.923
Treatment II	.812	.213	-.489	-.120	.919
valuable-worthless <sup>c</sup>					
Osgood <i>et al.</i>	.79	.04	.13	.00	.64
Treatment I	.948	.116	.040	-.145	.874
Treatment II	.744	-.089	.565	-.186	.838

TABLE 1 (continued)

Scales	I	II	Factors III	IV	$r^2$
sharp-dull					
Osgood <i>et al.</i>	.23	.07	.52	— .10	.34
Treatment I	— .205	.278	— .725	.561	.922
Treatment II	— .138	.059	— .202	.896	.750
strong-weak <sup>c</sup>					
Osgood <i>et al.</i>	.19	.62	.20	— .03	.46
Treatment I	.005	— .810	.519	.122	.886
Treatment II	.098	.880	.048	— .377	.861
clean-dirty					
Osgood <i>et al.</i>	.82	— .05	.03	.02	.68
Treatment I	— .915	— .203	— .082	.233	.884
Treatment II	— .729	.231	— .512	.237	.816
fast-slow					
Osgood <i>et al.</i>	.10	.00	.70	— .12	.50
Treatment I	.222	.415	— .803	— .052	.756
Treatment II	.413	— .621	.120	.520	.707
honest-dishonest					
Osgood <i>et al.</i>	.85	.07	— .02	.16	.75
Treatment I	— .943	— .067	— .167	— .012	.848
Treatment II	— .886	— .030	— .099	.066	.641
pleasant-unpleasant <sup>c</sup>					
Osgood <i>et al.</i>	.82	— .05	.28	— .12	.77
Treatment I	.947	.210	— .186	— .021	.954
Treatment II	.723	— .228	.605	.024	.887
active-passive <sup>c</sup>					
Osgood <i>et al.</i>	.14	.04	.59	— .02	.37
Treatment I	— .061	— .001	.961	.042	.863
Treatment II	— .215	.165	.439	— .772	.774
fragrant-foul <sup>c</sup>					
Osgood <i>et al.</i>	.84	— .04	— .11	.05	.72
Treatment I	.887	.374	— .085	— .028	.873
Treatment II	.433	— .547	.587	— .070	.700
large-small <sup>c</sup>					
Osgood <i>et al.</i>	.06	.62	.34	.04	.51
Treatment I	— .212	— .837	.215	.191	.686
Treatment II	.168	.899	— .197	— .032	.767
fair-unfair <sup>c</sup>					
Osgood <i>et al.</i>	.83	.08	— .07	.11	.71
Treatment I	— .843	.018	.365	.166	.759
Treatment II	— .929	.114	.022	— .207	.844
tasty-distasteful					
Osgood <i>et al.</i>	.77	.05	— .11	.00	.61
Treatment I	.959	.121	.104	.015	.893
Treatment II	.329	.049	.832	— .316	.872
hard-soft					
Osgood <i>et al.</i>	— .48	.55	.16	.21	.60
Treatment I	— .939	— .224	.070	.159	.926
Treatment II	— .897	.038	— .324	.120	.857

TABLE 1 (continued)

Scales	I	II	Factors III	IV	$h^2$
nice-awful					
Osgood <i>et al.</i>	.87	-.08	.19	.15	.82
Treatment I	.470	.870	-.038	.011	.959
Treatment II	.343	-.847	.320	.044	.881
beautiful-ugly <sup>c</sup>					
Osgood <i>et al.</i>	.86	.09	.01	.26	.82
Treatment I	-.953	-.265	.114	.055	.990
Treatment II	-.833	.274	-.442	-.118	.959
kind-cruel <sup>c</sup>					
Osgood <i>et al.</i>	.82	-.10	-.18	.13	.73
Treatment I	.848	.264	-.245	-.219	.804
Treatment II	.782	-.409	.331	-.085	.803
sacred-profane <sup>c</sup>					
Osgood <i>et al.</i>	.81	.02	-.10	.01	.67
Treatment I	-.220	-.497	.032	.810	.906
Treatment II	-.264	.775	.138	.430	.764
bitter-sweet <sup>c</sup>					
Osgood <i>et al.</i>	-.80	.11	.20	.03	.69
Treatment I	.935	.271	-.050	.008	.902
Treatment II	.380	-.201	.842	-.286	.953
Percent common variance					
Osgood <i>et al.</i>		Information not applicable			
Treatment I	63.39	17.05	8.74	4.20	
Treatment II	55.41	17.73	12.45	4.85	

Note: Directionality of the printed factor loading in the scientific subroutine employed is an arbitrary function of the first significant loading in each factor.

<sup>a</sup> Scales and factor loadings reprinted with permission of the publisher from C. E. Osgood, G. J. Suci, and P. E. Tannenbaum, *The Measurement of Meaning*. Urbana: Univ. Illinois Press, 1957. Table 1, page 37.

<sup>b</sup> Scale names are in the direction used by Osgood *et al.* except as otherwise indicated.

<sup>c</sup> For Treatments I and II, directionality of the scale was the reverse of that used by Osgood *et al.*

Comparison of Treatment II rotated factor loadings with those of Osgood *et al.* shows that 11 of the 21 major loadings agree with the published findings. Of the remaining major loadings four give rise to a new interpretation of Factor III, and two to an interpretable Factor IV (see Table 1).

The result of the Kolmogorov-Smirnov (8) two-sample test comparing the mean for each scale on each concept between Treatment I and Treatment II is  $z = 4.520$ ,  $p < < 3(10^{-8})$ .

### 1. Treatment I

With three exceptions, Factor I of Treatment I is in excellent agreement with the findings of Osgood *et al.* (5). The variance for the scale *hard-soft*

is located in Factor I instead of Factor II, indicating that at least for these *Ss* *hard-soft* has evaluative connotations. The variance for the *nice-awful* scale loaded in Factor II instead of Factor I. However, for this scale there is a significant minor loading (.470) in Factor I. The last of the differences is related to the *sacred-profane* scale. The variance for this scale is loaded primarily in the fourth factor (.810) with a significant minor loading (— .497) in the second factor. It is possible that these inconsistencies are a function of language differences, sex differences, and/or ethnic difference between the *S* populations employed by Osgood and his co-workers and those employed in this investigation. However, since Osgood and his co-workers do not give the necessary information, it would be difficult to determine the probable cause or causes of these shifts.

Factor II and Factor III are in nearly total agreement with the findings presented by Osgood and Suci (4) and Osgood *et al.* (5). Thus the results of Treatment I of this investigation represent a substantial replication of the original publication on which the investigation was based.

## 2. Treatment II

*Factor I.* The scales with major loadings in Factor I are good-bad, clean-dirty, honest-dishonest, unpleasant-pleasant, unfair-fair, hard-soft, ugly-beautiful, and kind-cruel; since most of these adjective pairs have connotations of morality in their normal usage, this factor has been called a *moral evaluative* dimension of meaning. Furthermore, the scales with significant minor factor loadings (.300 to .500) also have connotations of morality even though these connotations may not be their primary meanings.

*Factor II.* The three *potency* scales included in this investigation are in good agreement with Osgood, et al. (5) and Treatment I of this investigation. However, the additional significant minor factor loadings achieved in Treatment II appear to indicate greater sensitivity because sour-sweet, foul-fragrant, and cruel-kind have secondary connotations of potency in their normal usage in the English language.

*Factor III.* Factor III of Treatment II has a different interpretation from that in Treatment I, Osgood and Suci (4), or Osgood *et al.* (5). The major loadings are on the scales sour-sweet, foul-fragrant, tasty-distasteful, and bitter-sweet. The significant minor factor loadings (.300 to .600) are on the scales good-bad, worthless-valuable, clean-dirty, unpleasant-pleasant, passive-active, hard-soft, nice-awful, ugly-beautiful, and kind-cruel. The scales with major factor loadings and significant minor loadings in this factor have



important *esthetic* connotations. Factor III has been tentatively named *esthetic evaluative*, reflecting the implications of the findings.

*Factor IV.* With major loadings on the active-passive and sharp-dull scales, and significant minor loadings on the strong-weak and fast-slow scales, Factor IV is an adequate reflection of activity. This factor has been named *activity* in accordance with the name given Factor III in Treatment I of this investigation, Osgood and Suci (4), and Osgood *et al.* (5).

#### D. DISCUSSION

The modified instructions produce an instrument of greater sensitivity yielding finer nuance of meaning interpretable in logical terms from the different contextual connotations of the adjectives in English usage—both written and spoken. The adjectives heavy, light, good, bad, worthless, valuable, clean, dirty, honest, dishonest, pleasant, unpleasant, unfair fair, hard, soft, ugly, beautiful, cruel, and kind have connotations of *moral evaluation* in many of their contextual usages. Heavy, light, sweet, sour, weak, strong, fast, slow, foul, fragrant, small, large, nice, awful, cruel, kind, profane, and sacred have connotations of *potency* in their several contextual usages. The *esthetic evaluative* connotations of sour, sweet, bitter, good, bad, worthless, valuable, clean, dirty, pleasant, unpleasant, foul, fragrant, tasty, distasteful, hard, soft, nice, awful, ugly, beautiful, cruel, and kind are obvious. The *activity* connotations of sharp, dull, weak, strong, slow, fast, passive, active, fair, and unfair are generally recognized. Each of these adjectives has an activity or dynamic dimension to several of its connotations.

As discussed above, Treatment I results of this investigation are in good agreement with the findings of Osgood and Suci (4) and Osgood *et al.* (5); therefore the differences between the results of Treatment I Treatment II may be interpreted as a direct result of the modification to the middle interval instructions. The suggested modification of the instructions yields an instrument of greater sensitivity. For those investigators who desire or require this greater sensitivity the modified instructions may prove useful. For applications that do not require this greater sensitivity, the usual instructions of Osgood *et al.* (5) have proven their reliability over a great number of situations, concepts, and scales.

#### REFERENCES

1. DECHTER, N. Differences in self-conception and ideal role behavior among male and female Mexican-Americans and Anglo Americans of the Southwestern United States as measured by a semantic differential. Unpublished tutorial, Chatham College, Pittsburgh, Pennsylvania, 1966.

2. INTERNATIONAL BUSINESS MACHINES CORPORATION. System/360 Scientific Subroutine Package (360-CM-03X) Version III. White Plains, N.Y.: Internat. Business Machines, 1968.
3. OETTING, E. R. The effects of forcing response on semantic differential. *J. Educ. & Psychol. Meas.*, 1967, **27**, 699-702.
4. OSGOOD, C. E., & SUCI, G. J. Factor analysis of meaning. *J. Exper. Psychol.*, 1955, **50**, 325-338.
5. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. E. *The Measurement of Meaning*. Urbana: Univ. Illinois Press, 1957.
6. PLAKOVITZ, G. M. Secularization as change in concept meanings. Unpublished tutorial, Chatham College, Pittsburgh, Pennsylvania, 1964.
7. SHERIF, M., & SHERIF, C. W. *Social Psychology*. New York: Harper & Row, 1969.
8. SMIRNOV, N. Tables for estimating the goodness of fit for empirical distributions. *Ann. Math. Stat.*, 1948, **19**, 279-281.
9. SUCI, G. J. A comparison of semantic structures in American Southwest culture groups. *J. Abn. & Soc. Psychol.*, 1960, **61**, 25-30.
10. TRIANDIS, H. C., & OSGOOD, C. E. A comparative factorial analysis of semantic structures in monolingual Greek and American college students. *J. Abn. & Soc. Psychol.*, 1958, **57**, 187-196.

*Department of Philosophy and Psychology*  
*San Antonio College*  
*1300 San Pedro Avenue*  
*San Antonio, Texas 78212*

## MOTIVATIONAL DETERMINANTS OF FAMILY PLANNING CLINIC ATTENDANCE\*<sup>1</sup>

*The Behavioral Sciences Center, Bowman Gray School of Medicine,  
Wake Forest University*

CARL M. COCHRANE, CLARK E. VINCENT, C. ALLEN HANEY,  
AND ROBERT MICHIELUTTE

### SUMMARY

Southern, black, poverty-level, female subjects were classified as Actives, Dropouts, or Never-Beens in regard to attendance at a family-planning clinic. The subjects took a story-telling test designed for lower social class subjects and rated a variety of persons and activities along semantic differential dimensions. Women who had dropped out were significantly lower than the other groups in their needs for achievement and for controlling life events. Actives perceived more difficulty in controlling events. Several significant differences were also found in the semantic differential ratings. The results can be interpreted to yield a consistent set of hypotheses about motivations related to birth control decisions.

### A. INTRODUCTION

In view of the fact that population dynamics have become an area of major concern, it is surprising that research utilizing psychological concepts and methods in the area of family planning has not been abundant. This paucity of such research may be due in part to the lack of positive findings in early studies (2, 17, 18, 19) which attempted to relate psychological variables to fertility measures. Perhaps because of increasing sophistication in design and methodology, however, some positive results have recently begun to appear. These are described below.

The bulk of this work, which has been done in a variety of countries, has focused on the relationships between fatalistic *versus* activist attitudes and behaviors involving birth control. With use of a number of objective scales, several studies have measured feelings of being powerless in the face of

\* Received in the Editorial Office on February 12, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Supported in part by grants from the Ford Foundation and from the Welfare and Social Security Administrations, U.S. Department of HEW (CRD-283 and MC-R-370019).

external forces (7, 9, 20). Another made use of Seeman's (15, 16) constructs involving a sensed lack of internalized normative standards, and of the tendency to see events as chaotic and meaningless (3). Such feelings have been found to be related to larger ideal family estimates (7), higher birth rates (3), less favorable attitudes toward birth control (20), and lower frequencies of contraceptive use (9).

One factor analytic study of beliefs and attitudes related to contraceptive usage has been reported (1), and the same data reanalyzed to show that contraceptive users were more consistent in their attitudes and beliefs than non-users (6).

Family structure has also been subjected to research. Male dominance, restriction of the wife's role, and poor dyadic communication have been found to relate negatively to effectiveness and frequency of contraceptive practices (4, 5, 13).

The above studies used objective measurement techniques. In the only research using a projective approach, stories in children's reading books from a number of countries were scored in terms of the need for affiliation. Books from 1925 and 1950 were selected across countries; need for affiliation was correlated with birth rates in a *negative* direction during the early period and in a *positive* direction in the later period (10).

## B. METHOD

The study reported here made use of both projective and objective psychological measures in an attempt to ascertain motives related to utilization of a family-planning clinic by women in a sample of black, poverty-level subjects. This study was designed as an exploratory, or "pilot," effort to aid in selecting and perfecting instruments and in deriving hypotheses for further research.

### 1. *The Sample*

On the basis of 216 interviews completed in the subjects' homes, 166<sup>2</sup> women were selected for this study. The following criteria were applied: being a black female between the ages of 20 and 39 inclusively, having at least one living child, and being at a poverty level. [Poverty level was determined by use of a modification of Orshansky's Index (11). A minimum income of \$1900 was assumed necessary for a single individual, \$2700 for a couple, \$3200 for a three-person household, and approximately \$600 to \$700 for each additional member of the household.]

<sup>2</sup> Of the original 216 subjects, 15 were excluded because of contradictions as to clinic status, and 35 were dropped because of having incomes above poverty level.

On the basis of clinic records or reports of community-action workers, the women were classified into three groups: (a) Actives—those who had been active at the County Family-Planning Clinic for at least six months; (b) Dropouts—those who had dropped out of the Clinic; (c) Never-Beens—those who had never attended the Clinic despite an awareness of its existence. Of the 166 women in the pilot study, 51 were classified as Actives, 73 as Dropouts, and 42 as Never-Beens.

## 2. The Instruments

a. *The Need Evaluation Test.* The Need Evaluation Test (NET), constructed by the senior author, consists of nine pictures about which the subjects tell stories. The content of the pictures is divided into three situations or types of human relationships: adults with children, male and female, and a person confronting an authority figure. The costumes, situations, and details of the pictures were designed so that respondents from lower social classes could easily identify with them, and thus a variety of story responses to any picture would be possible.

The content of the stories was scored for the presence of four needs or drives, and four areas of perceived difficulties or obstacles, as follows:

(1). *Needs.* (a) Achievement—To succeed, achieve, do things well; (b) Affiliation—To have close, warm interpersonal relationships; (c) Control—To control events and other people so as to achieve needs or life goals; (d) Sex—To want sexual relationships in a pleasant, harmonious context.

(2). *Perceived difficulties.* (a) Control—People, events, or forces are seen as interfering with needs or life plans; (b) Authority—People in authority are resented as potentially imposing on or mistreating the subject; (c) Sex—Sexual situations are seen as negative in terms of satisfaction and fulfilling relationships; (d) Conflict—Situations are seen as involving interpersonal friction, ambivalences, unpleasant mood states, or other indications of unhappiness.

The difficulty of obtaining reliable scores on projective tests is well known. With this in mind, the test and the scoring procedures were planned in accordance with principles that were felt to be important in achieving reliable and valid scores.

Many projective tests furnish enough content detail to offer a variety of stimuli which interact with the subject's associative processes, thereby distracting him and offering multiple elements for him to interpret. The result is a complex story in which it is difficult to separate simple interpretations of details from more meaningful statements. Such complex results also lend



themselves to more intuitive judgments by the scorer, at the cost of less agreement between scorers.

To avoid such difficulties with the NET, the following steps were taken:

(a) The test pictures were drawn as simple line sketches. Details (face, hair, etc.) were not filled in, and clothing was portrayed as simply as possible (shirt and pants, or dress). This simplicity also tended to facilitate identification by lower-class subjects of either race with persons in the pictures. Other situational content was left out except for a few pieces of furniture.

(b) Scoring was done by laymen rather than by psychologists.

(c) It was emphasized to the scorers that they must judge *manifest content*, refraining from interpretation wherever possible.

(d) The scoring system was restricted to judgments that a theme was present (score 1) or absent (score 0) in a given story. This system is in accordance with McClelland's (10) finding that this method was as effective in scoring achievement needs as more intricate systems.

(e) All three scorers scored each subject's stories. If each scorer is considered to be a "test," then using three judges triples the length of the test and should increase the reliability of the scores.

b. *The Semantic Differential Scale.* The second instrument used was the Semantic Differential Scale, developed by Osgood, Suci, and Tannenbaum (12), who conducted extensive cross-cultural research in an effort to describe and measure attitudes (including emotional and affective reactions) toward people, events, and objects. They have reported evidence, based on factor analysis, that attitudes or affective judgments can be described basically as varying along three dimensions: Evaluation (good *versus* bad); Potency (big, strong *versus* little, weak); and Activity (active, vigorous *versus* inactive, passive).

The Semantic Differential Scale, designed to measure these three dimensions of affective judgments, includes a great many adjectives that are classified as to which dimension they measure. These adjectives can be arranged as polar opposites for a given dimension (e.g., "just" and "unjust" are at the opposite ends of the Evaluation dimension). The basic technique for measuring a subject's reaction to, or evaluation of, a concept is to present a set of such polar adjectives, sample each dimension, and ask the subject to indicate which adjectives best describe the concept. This is often done by placing the adjective pairs at each end of a line marked off with scale points. The subject can then indicate how close to either end he places the concept.

In this first study, we used a modified, three-step scale in order to make

the task as simple as possible. Four adjective pairs were given for each of the three dimensions (Evaluation, Potency, and Activity). The three-step scale was scored from 0 to 2 (e.g., on the "good-bad" pair, a rating of "good" would be scored as 2, an intermediate rating as 1, and a "bad" rating as 0). Scores on the adjective pairs were added to give a subject's total on that dimension, the range of possible scores being from zero to 8.

The concepts rated by our subjects were (a) having babies, (b) doctor, (c) the husband or current sexual partner,<sup>3</sup> (d) self, (e) men, (f) women, (g) policeman, and (h) alderman. It was felt that attitudes toward "having babies" and "doctor" relate directly to attitudes toward family planning, and that attitudes toward "self," "partner," "men," and "women" might relate dynamically to a woman's feelings about childbearing. "Policeman" and "alderman" were included to determine whether attitudes toward governmental and authority figures might be related to attendance at a county-sponsored clinic.

### 3. Scoring Reliability

To assess scoring reliability with the NET, intraclass correlation coefficients were derived from data on the first 45 subjects scored (15 from each clinic-status group). The intraclass correlation can be interpreted as representing the average of a set of correlations in which each scorer's ratings are correlated with each of the other two scorers across the subjects. This measure theoretically represents the reliability of a set of scores given by one judge. Coefficients representing the reliability of the scores when based on combined scores of the three judges were also computed (8). The reliability of the scoring dimensions, based on three judges, was adequate. Coefficients in the .80's were obtained for achievement, affiliation, need sex, and conflict, and in the .70's for control difficulties and sex difficulties. Agreement was lowest, but still reasonable for research purposes, on the need control and difficulty with authority scores (.50 and .63, respectively). The resentment of authority and the two sex variables had rather low frequencies of being scored as present. Underlying normality was assumed so as to compute the reliability data, but the more conservative approach of using chi squares was used in analyses relating to the experimental questions in the study.

### 4. Control Variables

The clinic-status groups did not differ significantly in regard to age, education, or marital status. The Never-Been group averaged significantly fewer

<sup>3</sup> Women living alone were asked to rate their most recent partner.

children than the other two groups, however. Consequently, the possible influence of parity on the psychological findings was assessed by further statistical analyses. No significant correlations between parity and any of the picture-story variables were obtained. An hypothesized correlation between parity and control difficulties was suggestive but not significant ( $R$  was .12, whereas .15 was needed for significance). Some of the semantic differential ratings were not normally distributed; but when the relations of parity to these ratings were examined by 24 chi squares (Evaluation, Potency, and Activity for each of the eight items rated), only one chi square was significant. This is quite probably a chance result.

### C. RESULTS

#### 1. *The NET*

Two-way analyses of variance, involving "clinic status" and "marital status," were performed with five of the NET variables (Table 1). "Resentment of authority" and the two sex variables were analyzed by chi-square techniques, comparing frequencies in each group obtaining scores of zero *versus* those giving some positive score.

When the category of *clinic status* was examined, the following results were obtained:

(a) Actives and Never-Beens were equal on the need for achievement and the need to control events and persons. Dropouts scored significantly lower than the Actives and Never-Beens on these two needs.

(b) Actives scored higher on control difficulties with people and events, Dropouts were intermediate, and Never-Beens scored lowest on this variable.

(c) Clinic status did not relate to the other NET variables.

When the category of *marital status* was examined, it was found to be related to only one variable. Women with no male partner perceived more difficulty in controlling their lives than women married or living with a partner. The means, not shown in Table 1, were as follows: Married, 3.7; Partner, 3.4; No partner, 5.0. Since neither clinic status nor marital status showed any significant differences for the variables analyzed by chi squares, the results are not tabled. Both authority difficulties and need sex were suggestively lower for the Never-Beens than for the other two groups, but the overall chi squares did not suggest that further partitioning would be of value.

#### 2. *Semantic Differential*

The ratings on the semantic differential dimension of Evaluation were decidedly skewed; the subjects tended to give positive ratings to all things

TABLE 1  
DIFFERENCES BETWEEN CLINIC-STATUS GROUPS ON THE NEED EVALUATION TEST (NET)

NET variable	Mean score by clinic status			<i>F</i> <sup>a</sup>	<i>p</i>
	Active ( <i>N</i> = 51)	Dropout ( <i>N</i> = 73)	Never- Been ( <i>N</i> = 42)		
Need achievement	3.6	2.4	3.6	3.2	<.05
Need affiliation	8.5	7.6	7.7	1.0	ns
Need control	13.2	10.4	13.4	4.1	<.001
Control difficulties	5.0	3.9	3.1	4.1	<.05
Conflict	6.0	6.4	6.1	1.0	ns

<sup>a</sup> Taken from two-way analyses of variance. *F* has 2 and 157 *df*.

rated. This was probably a function of three factors: (a) The subjects did feel positive about all the things rated. (b) The subjects may have been oriented toward pleasing the interviewer. (c) Our use of only a three-point scale (in an effort to make the task simple for subjects with low literacy levels) may have constricted differential evaluations. The statistical analyses were done by dividing the ratings into two categories: "more" favorable judgments included only the most favorable rating of 8; "less" favorable judgments included all other ratings. Chi squares were then performed to determine if the proportions of "more" favorable and "less" favorable ratings differed among the clinic-status groups. The results are shown in Table 2.

The clinic-status groups differed in their evaluations of four of the eight concepts: husband, men, alderman, and having babies. The direction of differences was the same in every case; the Never-Beens gave higher proportions of favorable ratings, while the Dropouts and Actives were roughly equal.

The groups did not differ on ratings of doctor, policeman, self, and women. Since all groups tended to rate doctor and self very highly, it might be argued that a "ceiling effect" was present, not allowing room for significant variation. Policeman and women received the lowest percentages of favorable evaluations in all clinic-status groups, however. The tendency of the Never-Beens to rate objects more favorably does not appear to be a set to give more positive evaluations of all items. The items they rated more favorably than the other two clinic-status groups seem to center on males (e.g., husband or partner, men, and alderman), with the exception, however, that a lower percentage of them gave a more favorable evaluation to policeman.

The intergroup differences in the ratings for husband or partner were greater (and the chi squares proportionally higher) than the differences in the ratings for men and alderman. The higher evaluation of husband or partner by Never-Beens presumably is consistent with their higher rating



of having babies, an activity necessitating a male. It is also possible that Never-Been women have a more optimistic view of men as a result of having a better relationship with their partner, which will be discussed later.

The Potency and Activity ratings were separated into "more" and "less" categories by division at the medians of the distributions. Chi squares were used to assess differences among the clinic-status groups. Although the Activity ratings were uniformly nonproductive of significant differences, several interesting trends were noted in the Potency ratings. These are shown in Table 2. When tested with an overall chi square, none of these trends was significant. However, all differences mentioned here are significant when a chi square is used comparing one activity group with the pooled remainder of the sample, or when order is specified in the statistical test: that is, when the chi square is a linear one stating that the value for group *a*, for example, should be greater than that for group *b*, which should be greater than that for group *c*. These results, while they are not conclusive, do merit interest.

TABLE 2  
PERCENTAGES<sup>a</sup> OF CLINIC-STATUS GROUPS GIVING "MORE FAVORABLE" OR  
"MORE POTENT" EVALUATIONS TO CONCEPTS RATED

	Clinic status				
Concept rated	Active ( <i>N</i> = 51)	Dropout ( <i>N</i> = 73)	Never- Been ( <i>N</i> = 42)	Chi square <sup>b</sup>	<i>p</i>
	<i>Evaluation</i>				
Having babies	43	49	76	11.4	<.01
Doctor	86	89	93	1.0	ns
Husband or partner	59	62	88	11.0	<.01
Self	86	84	88	<1.0	ns
Men	37	38	60	6.0	<.05
Women	69	74	71	<1.0	ns
Policeman	61	59	55	<1.0	ns
Alderman	72	67	88	6.2	<.05
	<i>Potency</i>				
Having babies	37	46	60	4.6	ns
Doctor	61	49	38	4.8	ns
Husband or partner	53	42	64	5.2	ns
Self	45	41	38	1.0	ns
Men	43	44	40	1.0	ns
Women	53	51	48	1.0	ns
Policeman	57	40	60	5.6	ns
Alderman	67	62	71	1.2	ns

<sup>a</sup> A percentage of 100 is possible for each entry.

<sup>b</sup> Based on contingency table with 2 *df*.



Table 2 shows that there is a linear relationship in regard to judgments concerning the potency of doctors. More potent ratings were given most often by Actives; least often by Never-Beens. Potency ratings on having babies reversed this order (Never-Beens saw having babies as potent most often; the Actives, the least). Dropouts differed from both other groups in seeing partners and policeman as "less" potent. Never-Beens gave the highest ratings for partner, while Actives gave intermediate ratings.

#### D. DISCUSSION

The studies described in the introduction found orientations toward activism and away from fatalism to be related to smaller ideal family-sizes, to favorable attitudes toward birth control, and to the more frequent use of contraception. The results reported here are more complex, and some of them appear to conflict with the previous studies. For example, subjects who have not attended clinics have need control scores which are equal to those of subjects who are attending; the Never-Beens also have control difficulty scores which are lower than those of Active subjects.

Three major differences between the present research and other studies are as follows: (a) the use of projective measures; (b) the fact that birth control was measured by the behavioral criterion of clinic attendance rather than by verbal or questionnaire methods; and (c) the restriction of the sample to poverty-level blacks, probably resulting in a sample of lower socioeconomic status than any of the other studies.

With these differences, any attempt to explain the conflicting results would be overly speculative. One possible exception is McClelland's (10) findings of relationships between birth rates and need affiliation scores across a number of countries, as contrasted with our study's lack of significant results in relation to this need. The McClelland technique was to correlate averages for whole countries. That such "ecological" correlations may totally misrepresent the relationships found at the level of individual prediction has been recognized by sociologists since Robinson's (14) paper on "ecological fallacy" was published. McClelland's findings thus may or may not hold true in regard to relationships between individual subjects' needs for affiliation and their birth control behaviors. In our sample, no such relationship was found.

The following interpretive conclusions can be drawn from the present study of urban, Southern, poverty-level black females.

Both the clinic Actives and the Never-Beens are higher than the Dropouts in the need for achievement and in the broader need to control one's life

progress. The Actives, however, perceive themselves as having difficulty in ordering and controlling their lives. The fact that the Actives tend to see their male partners as of intermediate potency and as "less good" suggests that this relationship is one key area of such difficulties. Family planning may then offer one avenue by which these subjects can successfully act to control and order events in their lives. In some cases, the wish to avoid having more children in a partnership of questionable stability may be a more specific motive.

The Never-Beens score equally high on the need for achievement and the need for control, but seemingly are more secure and untroubled than the Actives. They feel less difficulty in controlling events, and rate their partners as both "potent" and "good." Feeling more secure and optimistic about the future, they are more willing to expand their families. An indication that childbearing may be an avenue for acting out their needs to do something worthwhile (achievement) is contained in their tendencies to rate having babies as a "good" and "potent" activity.

Although scoring lower on the need to control and the need to achieve, the Dropouts do not rank themselves as "less good" or "less potent," which suggests that they are relatively unconscious of and/or untroubled by their lack of "positive" drives. Like the Actives, the Dropouts evaluate their husbands more unfavorably. They also give their husbands a lower rating on Potency than do the other two clinic-status groups; yet their perceived difficulties in control are not as high as those of the Actives. This fact suggests that their lack of need to strive against events renders them less troubled by deficiencies in their partner. We would speculate that they attended the clinic because of some combination of factors: number of children, discontent with partner, and/or contacts with an extensive action program in the city. Once having attended, however, their motivational structure was not such as to impel them to continue to attend.

The conclusions outlined above are tentative and are primarily offered as hypotheses which can be more specifically tested in future studies. The results suggest that there is potential value in investigating further (a) the links between the spouse or partner relationship and family planning and (b) means-ends relationships involving needs to achieve and control (the ends) and family-planning decisions (the means).

#### REFERENCES

1. BLAKE, R. R., INSKO, C. A., CIALDINI, R. B., & CHAIKIN, A. L. Beliefs and attitudes about contraception among the poor. Monograph 5, Carolina Population Center, University of North Carolina, Chapel Hill, 1969.

2. FLANAGAN, J. C. A study of factors determining family size in a selected professional group. *Genet. Psychol. Monog.*, 1942, **25**, 3-99.
3. GROAT, T. H., & NEAL, A. G. Social psychological correlates of urban fertility. *Amer. Sociolog. Rev.*, 1967, **32**, 945-959.
4. HILL, R. Putting the "family" in family planning. In Bogue, D. J. (Ed.), *Mass Communication and Motivation for Birth Control*. Chicago: Univ. Chicago Commun. & Fam. Study Center, 1967. Pp. 203-230.
5. HILL, R., STYCOS, J. M., & BACK, K. W. *The Family and Population Control: A Puerto Rican Experiment in Social Change*. Chapel Hill: Univ. North Carolina Press, 1959.
6. INSKO, C. A., BLAKE, R. R., CIALDINI, R. B., & MULAİK, S. A. Attitude toward birth control and cognitive consistency: Theoretical and practical implications of survey data. *J. Personal. & Soc. Psychol.*, 1970, **16**, 228-237.
7. KAHL, J. A. The measurement of modernism: A study of values in Brazil and Mexico. Latin American Monographs No. 12, University of Texas Press, Austin, 1968.
8. LINDQUIST, E. F. *Design and Analysis of Experiments in Psychology and Education*. Boston: Houghton Mifflin, 1956.
9. MACDONALD, A. P. Internal-external control and the practice of birth control. *Psychol. Rep.*, 1970, **27**, 206.
10. MCCLELLAND, D. C. *The Achieving Society*. Princeton, N.J.: Van Nostrand, 1961.
11. ORSHANSKY, M. Counting the poor: Another look at the poverty profile. *Soc. Security Bull.*, 1965, Jan., 3-29.
12. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The Measurement of Meaning*. Urbana: Univ. Illinois Press, 1957.
13. RAINWATER, L. *Family Design*. Chicago: Aldine, 1965.
14. ROBINSON, W. S. Ecological correlations and the behavior of individuals. *Amer. Sociolog. Rev.*, 1950, **15**, 351-357.
15. SEEMAN, M. On the meaning of alienation. *Amer. Sociolog. Rev.*, 1959, **24**, 783-791.
16. ———. The urban alienations: Some dubious theses from Marx to Marcuse. *J. Personal. & Soc. Psychol.*, 1971, **19**, 135-143.
17. WESTOFF, C. F., POTTER, R. G., & SAGI, P. C. *The Third Child: A Study in the Prediction of Fertility*. Princeton, N.J.: Princeton Univ. Press, 1963.
18. WESTOFF, C. F., POTTER, R. G., JR., SAGI, P. C., & MISHLER, E. G. *Family Growth in Metropolitan America*. Princeton, N.J.: Princeton Univ. Press, 1961.
19. WHELPTON, P. K., & KISER, C. V., Eds. *Social and Psychological Factors Affecting Fertility* (5 vols.). New York: Milbank Memorial Fund, 1946-1958.
20. WILLIAMSON, J. B. Subjective efficacy and ideal family size as predictors of favorability toward birth control. *Demography*, 1970, **7**, 329-339.

*The Behavioral Sciences Center*  
*Bowman Gray School of Medicine*  
*Wake Forest University*  
*Winston-Salem, North Carolina 27103*



## NAME STYLE AND CONSERVATISM\*

*University of Auckland, New Zealand*

---

ROGER BOSHIER<sup>1</sup>

---

### SUMMARY

Numerous name styles (e.g., John Smith, J. D. Smith) are used in Western society. It was suggested that name style may be a useful cue for predicting aspects of personality; in particular high scorers on the modified Wilson and Patterson (18) Conservatism scale were expected to prefer "conservative" rather than "liberal" name styles.

Across nine situations "J. D. Smith" was the most frequently preferred name style. Persons who most frequently used "John Smith" (i.e., their first and last name, rather than initials and their last name) were more liberal than persons using other name styles. Separate analyses of name style behavior in each of nine situations revealed that whilst "J. D. Smith" was preferred, respondents tended to vary their style to suit situations. Nevertheless in each situation adult education students, the most liberal of the three groups studied, were more inclined than members of the other two groups to employ "John Smith" as their name style. However, respondents who did not adapt their name style across situations (i.e., they used *one* name style in all situations) were significantly more conservative than persons who employed two name styles across the nine situations.

It was suggested that name style behavior of conservative persons indicates a lack of self-disclosure. Results were compatible with previous findings which show low self-esteem, sexual repression, superstition, race prejudice, and the need to remain anonymous to be concomitants of conservatism.

---

### A. INTRODUCTION

A person's name is a significant aspect of self. Most people feel their name is part of them, not fortuitous, but built in like an arm or finger. In some societies a child is not regarded as born until named, and in most parts of the world legal procedures are necessary before one can change a surname (4).

---

\* Received in the Editorial Office on February 13, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Lawrence Ferrie's assistance is acknowledged.



Despite the fact that every person has a name and rarely acts with neutrality if it is misspelled or mispronounced, the psychology of name behavior has received scant research attention. Several writers have studied unique or unusual names as a factor in social adjustment (2, 8, 11, 13), whilst Hartman (9) has speculated about the psychological functions served by an alias.

During studies by Boshier (3, 4) demonstrating that attitude to self is related to attitude towards one's name, it was observed that different name styles (e.g., John Smith, J. D. Smith, etc.) are habitually employed, and these vary from situation to situation.

Could name style be a useful cue for assessing personality? For instance, do persons who habitually disclose their first name differ significantly, in personality terms, from persons who habitually fail to disclose their first name? In this regard, Hartman (10) offered a subjective description of "possible personality implications of the different name styles. . . . *John J. Brown* . . . is the conventional American name style. Its use implies conformity to social customs. . . . A similar impression is suggested by the more simple form *John Brown*. The omission of the middle initial has, however, implications for greater frankness, directness, or even some individuality. . . . *J. J. Brown* . . . implies restraint or emotional constriction" (10, p. 289).

It is obvious that the context in which name styles are employed has to be recognized in any study of name behavior. Hartman argues, for example, that persons who use military titles no longer appropriate to their present position reveal a "need to demonstrate adequacy or to display past accomplishments" (10, p. 290). He also notes that "the easy transition to the use of first names upon short acquaintance is typical American business and social practice and reflects the prevailing approval of an extroverted type of adjustment. More inhibited or introverted persons continue to employ a formal style of address in their interpersonal contacts" (p. 293).

A large number of sexual, social, institutional, cultural, and psychological factors determine which name style is habitually used. Crucial variables would be the status of the sender/receiver, the employment of written or verbal means of communication, personality characteristics of the sender/receiver, and the context in which the name style is employed. In military service and school contexts name style use is clearly delineated. For example, conservative New Zealand schools, some modelled on the British public school system, require pupils to refer to one another by their surnames.

The restraint or emotional construction hypothesized by Hartman as

characterizing persons who habitually use initials rather than their first name is synonymous with Wilson's (17) description of the conservative personality. In the original validation of the Conservatism scale (18) high scorers were hypothesized to be religiously dogmatic, possess a right-wing political orientation, insist on strict rules and punishments, be ethnocentric, and prefer conventional art, clothing, and institutions. In a subsequent investigation conservative persons have been shown to be low in self-esteem (5). Elsewhere low esteem persons have been shown to dislike their own names (3, 4).

Thus a conservative person, almost by definition, should show a preference for conservative or "restrained" name styles (e.g., J. Smith, J. D. Smith). This study was therefore designed to ascertain which name style is most frequently employed in nine situations (e.g., "when writing a letter to the Prime Minister," "signing a check," etc.), the most frequently used name style *across* situations, sex differences in name style behavior, and the relationship between name style behavior and Conservatism scores. Regarding the latter aim, Boshier's (7) modification of the Conservatism scale, which has been shown to be reliable and valid, was used.

## B. METHOD

### 1. Subjects

Seventy subjects were drawn from three groups known on *a priori* grounds to be contrasting with respect to conservatism. Subjects were 26 adult education students (two men and 24 women) residing in Remuera, an upper class (16) suburb of Auckland city; 15 prison officers (14 men, one woman) from Waikune Goal; and 29 New Zealand army personnel (15 men, 14 women) employed at army headquarters, Auckland, New Zealand.

### 2. Procedure

Respondents completed a two page questionnaire. The first page contained the Conservatism (C) scale. The second page was headed as follows:

How do you write your name? People write their names in different ways to suit situations (e.g., John Smith, J. D. Smith, J. David Smith, John Smith, etc.). Write *your* name (or names) in the boxes below as you would when:

Writing a letter to the Prime Minister.

Altogether respondents were asked to indicate what name style was used in nine situations. These situations are listed in Table 3 in the Results section. Under each situation was a box wherein the respondent wrote his name.

## C. RESULTS

## 1. Construct Validity

Table 1, which presents mean Conservatism scale scores, adds weight to C scale construct validation material already published (17). Army respondents were slightly, though not significantly, more conservative than the prison officers ( $t = .29$ ), who were in turn significantly more conservative than the adult students ( $t = 5.50$ ,  $df = 39$ ,  $p < .0005$ ).

In most samples (6, 7, 18) conservatism increases monotonically with age. In each age group women are usually slightly more conservative than men. Note in the present sample 24 of the 26 adult students, who were the most liberal of the three groups, were women, whereas almost all the prison officers and just over half the army personnel were men.

The mean ages of the adult students were 50.34 years ( $SD = 9.27$ ), of the prison officers 36.13 years ( $SD = 9.81$ ), and of the army personnel 38.17 years ( $SD = 11.56$ ).

Because of the distribution of the sexes in each of the sample groups, women were significantly more liberal than men ( $t = 1.98$ ,  $p < .05$ ). Men had a mean C score of 50.80, women a mean C score of 43.56.

## 2. Preferred Name Styles

Name styles were coded according to the combination of initials and names presented in Table 2. Whilst each respondent wrote his own name, results are expressed by means of the surname "Smith," various initials, and the first name "John."

TABLE 1  
CONSERVATISM SCALE MEANS AND *SD*s FOR AGE AND SAMPLE GROUPS

Group	<i>n</i>	Mean	<i>SD</i>
60-69 years	4	36.50	6.96
50-59 years	18	48.33	15.16
40-49 years	20	49.10	14.91
30-39 years	15	41.93	15.33
20-29 years	13	49.76	8.02
Adult students			
Men	2	29.50	3.53
Women	24	35.54	11.57
Prison officers			
Men	14	42.35	16.45
Women	1	63.00	
Army personnel			
Men	15	51.60	10.06
Women	14	57.64	5.85

Across the nine situations "J. D. Smith" was the most frequently used name style, followed by "J. Smith," and "John Smith."

These data were derived by extraction of the most frequently used style employed by each respondent in the nine situations. Thus if a respondent used "J. Smith" in six situations and "John Smith" in the remaining three situations, his most frequently used style was "J. Smith." Thus 55.71 percent of respondents most frequently used "J. D. Smith," whilst 20 percent most frequently used "J. Smith," and 15.71 percent "John Smith."

Persons who most frequently used "John Smith" were significantly more liberal than persons who most frequently used either "J. Smith" ( $t = 1.97$ ,  $p < .05$ ) or "J. D. Smith" ( $t = 2.11$ ,  $p < .02$ ). There was no significant difference between the mean Conservatism scores of persons who used "J. . . ." or "J. D. . . ." as their most frequent name style ( $t = .55$ ).

### 3. Name Style and Conservatism Scores

Table 3 shows that respondents using "John Smith" when writing to the Prime Minister were significantly more liberal than respondents using "J. Smith" ( $t = 2.94$ ,  $df = 22$ ,  $p < .01$ ) or "J. D. Smith" ( $t = 3.81$ ,  $df = 43$ ,  $p < .001$ ) or "David J. Smith" ( $t = 2.36$ ,  $df = 22$ ,  $p < .05$ ).

There were no significant differences between the mean Conservatism scores of respondents employing differential name styles when "writing to your child's teacher," or when "signing a memo addressed to your employer," or when "signing a petition to be presented to your mayor." However, respondents employing "John Smith" were significantly more liberal than respon-

TABLE 2  
CONSERVATISM SCALE MEANS AND SDs FOR MOST FREQUENTLY USED STYLES AND FOR NUMBER OF NAME STYLES USED OVER NINE SITUATIONS

Name style	<i>n</i>	%	Mean	SD
<i>Most frequently used name style</i>				
John Smith	11	15.71	39.36	17.00
J. Smith	14	20.00	50.92	12.31
J. D. Smith	39	55.71	48.84	11.94
J. David Smith	1	1.43	9.00	
David J. Smith	3	4.29	46.00	13.11
Impossible to code	2	2.86	38.00	12.72
<i>Number of name styles used over nine situations</i>				
One (same name style used in all nine situations)	22	31.44	51.27	13.33
Two	32	45.70	44.43	14.75
Three	14	20.00	44.07	11.84
Four	2	2.86	53.50	16.26

TABLE 3  
CONSERVATISM SCALE MEANS AND SDs FOR PERSONS USING DIFFERENT  
NAME STYLES IN NINE SITUATIONS

Name style	N	Prime Minister		N	Teacher		N	Employer		N	Petition	
		M	SD		M	SD		M	SD		M	SD
John Smith	13	38.3	12.18	22	40.86	15.46	2	45.00	5.65	10	42.50	18.81
J. Smith	11	52.00	11.29	11	52.90	13.36	14	52.64	11.01	17	47.76	12.12
J. D. Smith	32	52.75	11.27	23	51.04	12.17	29	51.27	10.88	33	49.36	12.30
J. David Smith				2	44.50	6.36	No subjects			One subject		
David J. Smith	11	49.00	9.56	4	48.25	6.84	3	52.66	11.93	7	46.14	11.28
Impossible to code	2	38.00	12.72	3	38.00	9.00	4	44.75	14.24	2	43.50	4.94
Non applicable	1			5			18*					

dents employing "J. D. Smith" ( $t = 3.03$ ,  $df = 42$ ,  $p < .01$ ) or "J. Smith" ( $t = 2.29$ ,  $df = 26$ ,  $p < .05$ ) when "appending your name to a notice to be circulated in your neighborhood."

There were no significant differences between mean Conservatism scores of respondents employing differential name styles when "writing to your local council . . . ."

Respondents "signing a check" with their first and last name—"John Smith"—were significantly more liberal than respondents "signing a check" with two initials and their surname—"J. D. Smith" ( $t = 2.09$ ,  $df = 44$ ,  $p < .02$ ). Similarly, respondents using their first and last names when "writing a letter to a newspaper" were significantly more liberal than persons using "J. D. Smith" ( $t = 2.69$ ,  $df = 44$ ,  $p < .05$ ).

Respondents using "John Smith" when "applying for a job" were more liberal than respondents using "J. Smith" or "J. D. Smith" or other name styles, although in these cases the differences were not statistically significant.

Summation within name styles and sample groups across the nine situations (see Table 4) reveals that the "liberal" adult education participants were significantly more inclined to use "John Smith" than other name styles (collapsed) than were the "conservative" army sample ( $\chi^2 = 68.07$ ,  $df = 1$ ,  $p < .001$ ) or the "conservative" prison officer sample ( $\chi^2 = 24.35$ ,  $df = 1$ ,  $p < .001$ ).

Coupled with results demonstrating the relationship between individual C score and name style, these data strongly suggest that conservative/liberal people habitually use different name styles. Differences identified did not occur as a function of sex, since men and women did not employ different name styles ( $\chi^2 = 1.30$ ,  $df = 4$ , ns).



TABLE 3 (continued)  
CONSERVATISM SCALE MEANS AND SDs FOR PERSONS USING DIFFERENT  
NAME STYLES IN NINE SITUATIONS

Notice			Council			Check			Newspaper			Job		
n	M	SD	n	M	SD	n	M	SD	n	M	SD	n	M	SD
13	36.92	16.08	7	42.00	17.16	8	37.50	16.80	14	38.35	18.67	16	46.06	12.30
15	49.66	13.29	15	49.73	13.94	14	48.85	14.18	15	47.13	14.93	10	49.60	14.63
31	50.51	12.40	42	47.16	12.58	38	48.52	12.81	32	50.18	10.99	26	50.84	11.47
One subject			One subject			3 32.66 20.98			One subject			2 21.50 17.67		
7	46.71	12.65	2	53.00	7.07	6	50.16	7.83	4	51.25	8.30	12	49.66	9.87
3	38.66	9.07	2	38.00	12.72	One subject			2	38.00	12.72	2	38.00	12.72
			1						2			2		

Note: Definitions of abbreviations used in column headings for the nine situations are as follows: Prime Minister—Writing a letter to the Prime Minister; Teacher—Writing to your child's teacher; Employer—Signing a memo addressed to your employer; Petition—Signing a petition to be presented to your mayor; Notice—Appending your name to a notice to be circulated in your neighborhood; Council—Writing to your local council complaining about rubbish collections; Check—Signing a check; Newspaper—Writing a letter to a newspaper; Job—Applying for a job.

\* These were housewives who had no employer.

#### 4. Adapting Names Style To Suit Situations

Table 2 showed in the nine situations that the most frequently used number of name styles was two. However, two of the 70 respondents used four styles in the nine situations. Respondents using only one name style in the nine situations were significantly more conservative than persons using two ( $t = 1.74$ ,  $df = 52$ ,  $p < .05$ ) or three name styles ( $t = 1.64$ ,  $df = 34$ ,  $p < .05$ ).

#### D. DISCUSSION

The above findings may be regarded as revealing a lack of self-disclosure (12) among conservative respondents. According to some writers self-disclosure is an identifying criterion of the healthy personality. "Third force"

TABLE 4  
NAME STYLE MOST FREQUENTLY USED ACROSS NINE SITUATIONS BY THREE SAMPLE GROUPS

Name style	Adult students		Army personnel		Prison officers	
	n	%	n	%	n	%
John Smith	74	36.45	14	5.74	17	12.69
J. Smith	26	12.81	50	20.50	46	34.33
J. D. Smith	79	38.92	149	61.10	58	43.29
J. David Smith	9	4.43	1	.41	2	1.49
David J. Smith	15	7.39	30	12.30	11	8.20
Total	203	100	244	100	134	100

psychologists (e.g., Fromm, Maslow, Rogers) have variously labelled the tendency to misrepresent oneself or the habitual presentation of only one part of oneself as a characteristic of the marketing personality, other-directedness, self-alienation, or as an indicant of a lack of full functioning or self-actualization. Certainly, the habitual nonuse of one's first name seems compatible with a lack of authenticity, self-rejection, and insecurity. Indeed, according to "third force" psychologists, neurosis is related to an ability to know oneself but also to an inability to make it known to others.

Conservatism is not necessarily pathological. Nevertheless, there is now considerable evidence (16) to suggest that it is a unitary syndrome serving psychological functions, rather than a cluster of social attitudes based on some objective evaluation of facts relating to social issues. The notion that high C score people lack self-disclosure is compatible with previous findings that demonstrate conservatism to be a general factor (6) having as concomitants low-self esteem (5), sexual repression (15), race prejudice (1), and the need to remain anonymous (14).

#### REFERENCES

1. BAGLEY, C. Racial prejudice and the conservative personality. *Polit. Stud.*, 1970, **18**(1), 134-141.
2. BAGLEY, C., & EVAN-WONG, L. Psychiatric disorder and adult and peer group rejection of the child's name. *J. Child Psychol. & Psychiat.*, 1970, **11**, 19-27.
3. BOSHER, R. W. Self-esteem and first names in children. *Psychol. Rep.*, 1968, **22**, 762.
4. ———. Attitudes towards self and one's proper names. *J. Individ. Psychol.*, 1968, **24**, 63-66.
5. ———. A study of the relationship between self-concept and conservatism. *J. Soc. Psychol.*, 1969, **77**, 139-140.
6. ———. To rotate or not to rotate: The Question of the Conservatism scale. *Brit. J. Soc. & Clin. Psychol.*, 1972, **11**, 313-323.
7. ———. An empirical investigation of the relationship between conservatism and superstition. *Brit. J. Soc. & Clin. Psychol.*, in press.
8. ELLIS, A., & BEECHLEY, R. M. Emotional disturbance in children with peculiar given names. *J. Genet. Psychol.*, 1954, **85**, 337-339.
9. HARTMAN, A. A. Criminal aliases: A psychological study. *J. of Psychol.*, 1951, **32**, 49-56.
10. ———. Name styles in relation to personality. *J. Gen. Psychol.*, 1958, **59**, 289-294.
11. HARTMAN, A. A., NICOLAY, R. C., & HURLEY, J. Unique personal names as a social adjustment factor. *J. Soc. Psychol.*, 1968, **75**, 107-110.
12. JOURARD, S. The transparent self. Princeton, N.J.: Van Nostrand, 1964.
13. McDAVID, J., & HARARI, H. Social desirability of names and popularity in grade school children. Paper given to A.P.A. convention, Chicago, 1965, cited by Hartman *et al.* (11).
14. PATTERSON, J. R., & WILSON, G. D. Anonymity, occupation, and conservatism. *J. Soc. Psychol.*, 1969, **78**, 263-266.
15. THOMAS, D. R., SHEA, J. D., & RIGBY, R. G. Conservatism and response to sexual humour. *Brit. J. Soc. & Clin. Psychol.*, 1971, **10**, 185-186.

16. TIMMS, D. W. G. *The Urban Mosaic*. London: Cambridge Univ. Press, 1971.
17. WILSON, G. D., *Ed.* *The Psychology of Conservatism*. London: Academic Press, 1973.
18. WILSON, G. D., & PATTERSON, J. R. A new measure of Conservatism, *Brit. J. Soc. & Clin. Psychol.*, 1968, **7**, 264-269.

*University of Auckland*

*Private Bag*

*Auckland*

*New Zealand*



## POLARITY AND "ACCURACY" OF RATINGS AND THE MEANINGFULNESS OF PERSONALITY DIMENSIONS\*<sup>1</sup>

*University of California, Berkeley*

---

WALTER C. BORMAN AND WILLIAM K. GRAHAM

---

### SUMMARY

This study investigated the relative polarity and "accuracy" of ratings on personally meaningful versus meaningless personality dimensions. In contrast to previous studies, ratings on dimensions regarded as meaningful were no more polar than ratings on dimensions regarded as meaningless. With respect to accuracy, ratings on the *least* meaningful personality dimensions were significantly more accurate than ratings on meaningful dimensions when personality test scores and self-ratings were used as criteria. When mean peer ratings were used as criteria, accuracy was higher regardless of the meaningfulness of the dimensions. The results were interpreted in terms of the kinds of dimensions selected as most and least meaningful.

---

### A. INTRODUCTION

The present study deals with two research questions stimulated by D. A. Kelly's (11) Personal Construct Theory: (a) Will ratings have greater polarity on dimensions chosen by subjects as most personally meaningful? (b) Are judges more "accurate" in their ratings on those dimensions most meaningful to them? Kelly proposed that a person's own construction of the phenomenal world is more meaningful to him than constructions imposed by others. This basic idea has stimulated a considerable amount of empirical research in the area of interpersonal perception. In some of these studies the meaningfulness of dimensions has been inferred from responses to Kelly's Repertory Grid Test (4, 9, 10, 12). In other studies subjects have been asked to select personally meaningful dimensions from a provided list (13, 14).

#### 1. *Polarity of Ratings*

Most studies of polarity have concluded that when raters are permitted to use personally meaningful dimensions, they tend to use the extremes of the

---

\* Received in the Editorial Office on February 20, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Reprint requests should be sent to William K. Graham at the address shown at the end of this article.



ratings scale to a greater extent than when required to rate on dimensions that lack personal meaning. Recent reviews by O'Donovan (14, 15) and Hamilton (18) suggest that individual characteristics of raters are also related to the polarity of ratings. O'Donovan, for example, found that raters with neurotic tendencies are more likely to use the extremes of a rating scale than "normal" raters. More recently, Bonarious (1) has proposed a model in which J (judge or rater), C (construct or dimension), and O (object or ratee) all have an effect on polarity. In a series of experiments, evidence was found for C and O effects and  $C \times O$  interactions. Little or no J effect was found when judges were placed on a normal-neuroticism continuum according to their scores on a Dutch personality test.

## 2. Accuracy of Ratings

The problems of ascertaining the accuracy of judges' perceptions of others have been well documented (e.g., 2, 3, 6, 16). First, there is the problem of what kinds of rating situations are best for obtaining valid indices of interpersonal accuracy. The variance in accuracy controlled by situational factors is probably considerable. Rating situations investigated in studies of accuracy have included (a) judging acquaintances, (b) rating attributes of characters in a movie, and (c) judging responses or characteristics of people just met in the experimental situation.

Second, there is the measurement problem explicated by Cronbach (5). Methods of obtaining accuracy scores from interpersonal perception experiments can vary greatly. Cronbach suggested several measures of accuracy that could be extracted from data in these studies. Cline (3) provided empirical evidence of relatively low intercorrelations among some of the measures; hence, the problem of which measure to choose may be of practical importance. Given the variety of conditions under which interpersonal accuracy has been and could be studied, and the variety of accuracy measures, consistency of findings across studies can hardly be expected.

Nonetheless, it should be possible to decide, on a *logical* basis, which accuracy measure or measures to use in a given situation. For example, if one were interested in ascertaining the accuracy of a curriculum evaluator's attempts to rank order the abilities of a given class in a number of areas, such as mathematics and reading comprehension, Cronbach's Stereotype Accuracy correlation index would be appropriate. This measure is a correlation between the evaluator's judgments of class members' overall ability level in the various areas and the actual mean scores for the class in these areas. If another type of accuracy were of interest, a different accuracy measure might

be utilized. For the present investigation, Cline's (3) Interpersonal Accuracy measure was chosen. This particular index was selected for two reasons. First, the type of accuracy of interest was the raters' ability to differentiate among ratees on each trait. Second, as Cline has pointed out, the Interpersonal Accuracy index is not affected by different scoring methods.

## B. METHOD

Fourteen college sorority girls were asked to rate themselves and one another on 10 personality dimensions defined by the Guilford-Zimmerman Temperament Survey: General Activity, Restraint, Ascendancy, Sociability, Emotional Stability, Objectivity, Friendliness, Thoughtfulness, Personal Relations, and Masculinity. The ratings were recorded on nine-point scales (low to high for each trait). In addition, subjects took the Guilford-Zimmerman Test. Finally, each subject indicated which three dimensions were most meaningful to her in thinking about and judging people in general, and which three dimensions were least meaningful in the same sense.

Polarity of ratings for a given judge on a given dimension was determined by subtracting each rating from the midpoint (5.0) and summing the absolute values over the 13 ratees. The polarity score for each judge's three most meaningful dimensions was simply the sum of that judge's polarity scores for those three dimensions. Polarity on least meaningful dimensions was calculated the same way. The mean of the 14 judges' polarity scores for most meaningful dimensions was then compared to the mean polarity for least meaningful dimensions.

Accuracy for a given judge on a given dimension was determined by correlating the judge's ratings of the 13 others on that dimension with the others' (a) self-ratings on that dimension, (b) personality test scores on that dimension, and (c) mean peer ratings on that dimension. Thus, for each judge there were three separate accuracy scores for each of the 10 dimensions. For accuracy in terms of self-ratings, there were 10 correlations for each judge, one for each of the 10 dimensions. Similarly, for each judge there were 10 correlations for the mean peer ratings and 10 correlations for the personality test scores.

The hypothesis that judges are more accurate on dimensions they regard as most meaningful to them was tested for each of the three accuracy criteria. Fisher  $r$  to  $z$  transformations were performed on the accuracy correlations corresponding to the three most and three least meaningful dimensions. For each judge and for each criterion, a mean was computed for the three  $z$  scores corresponding to most meaningful dimensions. The same procedure was fol-

lowed for the three least meaningful dimensions. For each criterion the mean of the  $z$  scores for most meaningful dimensions was compared to the mean of  $z$  scores for the least meaningful dimensions.

### C. RESULTS

There was considerable variability in the choices of most and least meaningful personality dimensions with the exception of Thoughtfulness and Friendliness which appeared consistently on judges' most meaningful lists (see Table 1).

In contrast to previous studies, the mean polarity for personally most meaningful dimensions was not found to be greater than polarity for the least meaningful dimensions. The mean distance of ratings from the midpoint of the scale was 1.64 for most meaningful and 1.67 for least meaningful dimensions.

Accuracy was found to be significantly greater on *least* meaningful dimensions as compared to most meaningful dimensions when personality test scores and self-ratings were used as criteria (see Table 2). When mean peer ratings were used as criteria, accuracy on *most* meaningful dimensions was higher but not significantly so.

### D. DISCUSSION

The polarity hypothesis as formed by O'Donovan (14) and Bonarious (1) was not confirmed here when judges were restricted to choosing personally meaningful dimensions from the 10 personality traits corresponding to the Guilford-Zimmerman. It is possible, of course, that the range of meaningfulness for judges across these 10 dimensions was not large enough to make any difference in the polarity of ratings.

The hypothesis that judges would be more accurate on dimensions most personally meaningful to them was not confirmed. The variances of ratings

TABLE 1  
FREQUENCY OF PERSONALITY DIMENSION CHOICES

Guilford-Zimmerman dimensions	One of three most meaningful	One of three least meaningful
1. General Activity	2	5
2. Restraint	2	6
3. Ascendancy	2	5
4. Sociability	1	7
5. Emotional Stability	6	0
6. Objectivity	1	7
7. Friendliness	11	0
8. Thoughtfulness	12	0
9. Personal Relations	4	3
10. Masculinity	1	9

TABLE 2  
MEAN INTERPERSONAL ACCURACY CORRELATIONS ( $N = 14$ ) FOR MOST AND  
LEAST MEANINGFUL DIMENSIONS

Criteria	Three <i>most</i> meaningful	Three <i>least</i> meaningful	$t^a$
Personality test scores	.18	.42	2.47, $p < .05$
Mean peer rating	.71	.68	.58(n.s.)
Self-rating	.19	.39	2.86, $p < .01$

<sup>a</sup> Tests are two tailed, 26 *df*.

and criteria were checked for possible restriction of range. Since there was less variability on criteria for most meaningful dimensions, corrections for restriction of range were applied (7). As a result, the accuracy correlation for personality test criteria on the meaningful dimensions increased from .18 to .21. For the self-rating criteria, the correlation for the most meaningful dimensions was raised from .19 to .25. Neither of these increases came close to bringing the accuracy indices to the level obtained for least meaningful dimensions. Hence, the greater accuracy on least meaningful dimensions cannot be attributed to restriction of range.

An alternative explanation for the obtained results is that ratings on Friendliness and Thoughtfulness—the dimensions most frequently selected as being meaningful—are peculiar in the sense that they tend to be based on personal relationships rather than on observations of the ratees' interactions with others. On the other hand, such dimensions as Masculinity and Objectivity—the dimensions most frequently selected as least meaningful—may be less subject to idiosyncratic rater-ratee relationships. This phenomenon could lead to consistently higher accuracy correlations on dimensions less affected by individual relationships.

To summarize, it was expected that ratings would be more polarized on dimensions personally meaningful to raters. It was also expected that the greater ease in differentiating people on constructs with which the raters were comfortable and familiar would lead to greater accuracy. That these hypotheses were not supported may have been due to the limited choice of dimensions, to the peculiar nature of the dimensions selected as meaningful, or both. In any case, it has been demonstrated that under some circumstances, the accuracy of ratings may be greater for relatively meaningless as opposed to meaningful dimensions of personality.

#### REFERENCES

1. BONARIOUS, J. C. J. Personal construct psychology and extreme response style: An interaction model of meaningfulness, maladjustment, and communication. Unpublished monograph, University of Groningen, Groningen, Netherlands, 1970.



2. BRUNER, J. S., & TAGIURI, R. The perception of people. In G. Lindzey (Ed.), *Handbook of Social Psychol.* Cambridge, Mass.: Addison-Wesley, 1954. Pp. 634-654.
3. CLINE, V. B. Interpersonal perception. In B. A. Maher (Ed.), *Progress in Experimental Personality Research (Vol. 1)*. New York: Academic Press, 1964. Pp. 221-284.
4. CROMWELL, R. L., & CALDWELL, D. F. A comparison of ratings based on personal constructs of self and others. *J. Clin. Psychol.*, 1962, **18**, 43-46.
5. CRONBACH, L. J. Processes affecting scores on "understanding of others" and "assumed similarity." *Psychol. Bull.*, 1955, **52**, 177-193.
6. GAGE, N. L., & CRONBACH, L. J. Conceptual and methodological problems in interpersonal perception. *Psychol. Rev.*, 1955, **62**, 411-422.
7. GHISELLI, E. E. *Theory of Psychological Measurement*. New York: McGraw-Hill, 1964.
8. HAMILTON, D. L. Personality attributes associated with extreme response style. *Psychol. Bull.*, 1968, **69**, 192-203.
9. ISAACSON, G. A comparative study of the meaningfulness of personal and cultural constructs. Unpublished Master's thesis, University of Missouri, Columbia, 1962, cited in O'Donovan, D., Rating extremity: Pathology or meaningfulness? *Psychol. Rev.*, 1965, **72**, 358-372.
10. ISAACSON, G. S., & LANDFIELD, A. W. Meaningfulness of personal versus common constructs. *J. Individ. Psychol.*, 1965, **21**, 160-166.
11. KELLY, G. A. *The Psychology of Personal Constructs*. New York: Norton, 1955.
12. LANDFIELD, A. W. Meaningfulness of self, ideal and other as related to own vs. therapists's personal construct dimensions. *Psychol. Rep.*, 1965, **16**, 605-608.
13. MITSOS, S. B. Personal constructs and the semantic differential. *J. Abn. & Soc. Psychol.*, 1961, **62**, 433-434.
14. O'DONOVAN, D. Polarization and meaningfulness in 6300 value judgements. Unpublished manuscript, University of Missouri, Columbia, 1964, cited in O'Donovan, D., Rating extremity: Pathology or meaningfulness? *Psychol. Rev.*, 1965, **72**, 358-372.
15. O'DONOVAN, D. Rating extremity: Pathology or meaningfulness? *Psychol. Rev.*, 1965, **72**, 358-372.
16. TAGIURI, R. Person perception. In G. Lindzey & Aronson (Eds.), *Handbook of Social Psychology (Vol. 3)*. Reading, Mass.: Addison-Wesley, 1969. Pp. 395-450.

*Department of Psychology*  
*University of California, Berkeley*  
*Berkeley, California 94720*



## PREMACKIAN REINFORCEMENT OF CLASSROOM BEHAVIOR THROUGH TOPIC SEQUENCING\*

*Arizona State University*

---

JACK C. HARTJE

---

### SUMMARY

In a test of Premack's model of reinforcement, 108 fifth-grade students at Cactus Wren Elementary School were randomly assigned to three groups. All subjects participated in 15 daily 90-minute experimental sessions. Prior to initiating the treatment, all of the subject's preferences for reading, social studies, and science were measured by means of a 10-point rating scale. Ratings for each of the three topics ranged from zero ("I don't like it") to 10 ("I like it very much"). On the basis of the pretest ratings, it was possible to determine for each subject his least preferred (L), intermediate or middle (M), and highest rated (H) topic.

One of the treatment groups, designated Reward, studied the three topics in a sequence that conformed to Premack's prescription for reinforcement. These subjects were required to achieve some criterion level of performance in L in order to gain access to M. Criterion achievement in topic M earned the right to initiate activities related to the most preferred topic, H. Posttest preference measures indicated that Reward preference ratings for L were significantly higher than pretest ratings ( $p < .01$ ).

Premack also proposed that punishment is essentially the reverse of reinforcement. The individual is forced to terminate a relatively high strength activity in order to initiate a relatively low strength behavior. In the present study, this notion was tested by requiring Punishment subjects to begin each day's session studying H. After 30 minutes, these subjects were required to move on to their intermediate topic, M. After another 30 minutes, they were required to study their least preferred topic, L. It was predicted and confirmed ( $p < .01$ ) that Punishment posttest preference measures of H would show a significant decrement.

Control group subjects were allowed to choose freely which of the three topics they would study. This constituted the only distinction between the

---

\* Received in the Editorial Office on February 20, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

two treatment groups and the Control group. It was predicted that since Control subjects did not undergo either a reinforcing or a punishing Premackian contingency, posttest preference measures would not differ from pretest values. Unfortunately, Control scores changed in the same manner as the two treatment groups. Several reasons why the treatment effect occurred for Control subjects were discussed.

---

### A. INTRODUCTION

Education can reasonably be described as the process of transferring information from a source to a recipient. The newly acquired information may serve to create behavioral capabilities or strengthen already existing ones. The information is typically categorized into various homogeneous topics, such as science, social studies, or reading skills. Considerable attention has been directed toward how best to present a given topic. Such factors as mode of presentation, feedback of results, and sequencing within a topic all have been experimentally demonstrated to be relevant variables.

One of the variables proposed to be of importance to learning and, therefore, worthy of control involves the sequence in which a particular topic is presented (7). Material that is presented with an emphasis upon the sequence or order of presentation is referred to as "programmed." There are a variety of definitions of programmed instruction. Dolmatch (1) defined a program as "a sequence of items, steps, or frames which present material to the learner. Each frame contains new information and/or a recapitulation of information, combined with some material requiring a response" (p. 13). Glaser (2) characterized programming as "the process of constructing sequences of instructional material in a way that maximizes the rate of acquisition and retention, and enhances the motivation of the student" (p. 7). Interest in programming is due in large part to a current trend favoring individualized instruction, for which programmed instruction is well suited. The substantial interest in determining how best to sequence material within a given topic has failed to generate a parallel concern for how best to sequence the topics themselves. Is there a rationale that would enable the educator to select a sequence of topics that would enhance learning?

It is not unreasonable to propose that education completely lacks a generally accepted theoretical basis to justify a particular ordering of topics. There is no reason to suspect that teaching science *after* social studies is superior (in terms of learning science) to the reverse ordering.

Premack's model of reinforcement (3) suggests a procedure by which

appropriate topic sequencing can be objectively determined. If the performance of a given low probability response (L) is followed by the opportunity to perform a more preferred response (H), L will become more probable. Generalizing to the sequencing of topics, if science is preferred to social studies, the requiring of some criterion behavior in social studies prior to the study of science should serve to reinforce behavior relevant to social studies. If reading is preferred to science (and thus to social studies), a similar contingency between science and reading should serve to reinforce behavior relevant to science, etc.

Premack's (5) interpretation of punishment proposes that since a transition from a less to a more probable event *facilitates* responding, a transition from a more to a less probable event may *suppress* responding. Premack completed several animal studies which support his analysis of punishment (4).

In her doctoral dissertation (8), Wasik made several tests of the applicability of Premack's principle to lever-pressing behavior. One of these tests involved multiple contingencies. Thirty-four male and female college students participated in four experiments. Various precontingency response probabilities were established for several levers by associating each lever with a different reinforcement schedule (CRF, FR5, FR10, and FR20).

In the first experiment involving two levers, Premack was supported. Access to the more probable lever served to reinforce pressing the less probable lever, but the reverse contingency had no reinforcing effect.

The second experiment also involved two levers associated with different probabilities. A period of time during which access to the more preferred lever was contingent upon first pressing the less preferred lever (contingency) was followed by a free access period (no contingency). In successive alternations of these two conditions, it was found that the reinforcing effects obtained during the contingency did not carry over into the free access period.

In the third experiment, contingencies were arranged for all possible pairs of the four levers. The most probable lever reinforced all others; the least probable reinforced no others. The two intermediate probability levers reinforced less probable and failed to reinforce more probable responses.

In the discussion so far, all of Wasik's experiments have supported Premack through the use of response pairs. In a test involving multiple contingencies, Premack was not supported. College students completed a contingency in which three responses in a chain were ordered from least to most preferred, consistent with Premack's principle. Inconsistent with Premack was the finding that subjects also completed a response chain in which

a less preferred response stood procedurally in the position of reinforcing a more preferred response, although the terminal member of the chain was the most preferred response. On the basis of this last experiment, Wasik concluded that a revision in Premack's basic generalizations is needed if the generalizations are to apply to response chains. Premack's current generalizations are seen by Wasik as applying only to paired responses.

In consideration of the fact that the terminal member of the response chain was the most probable member of the set of responses, Wasik's conclusions are open to question. Wasik found that a M, L, H sequence was "successfully completed." She concluded that L reinforced M and thus contradicted Premack. Another possibility is that H indirectly reinforced M with L simply serving to delay access to H. Since the present study involves a three-response chain, additional light will be shed on this discrepancy.

Although initially tested in animal reinforcement situations, the Premack principle has recently been applied in a variety of human settings. Wasik (9) carried out a test of the effectiveness of reinforcement outside the contingency with lever-pressing behavior using 12 college students. When a more probable response was made contingent upon a less probable response, there was an increase in strength for the less probable response. This reinforcing effect, however, did not carry over to subsequent postcontingency conditions.

Wasik (10) successfully employed a behavior management procedure in a demonstration school for culturally deprived children to increase appropriate behavior in a second-grade classroom. A free-choice activity time was introduced into the classroom, and access to this activity was made contingent upon the prior occurrence of desirable behavior. The average percentage of desirable classroom behavior increased during the contingency session, decreased when the contingency was removed, and increased again during the reintroduction of the contingency.

The present study tested Premack's principle in an applied classroom setting. Both the reinforcement and the punishment aspects were tested. If Premack's model can be extended to the classroom, an effective and relatively simple reinforcement technique will have been demonstrated. Not only will the external validity of the theory be strengthened, but applied classroom control will be enhanced.

## B. METHOD

### 1. *Questions Tested*

The purpose of the study was to determine the applicability of the Premack principle to classroom behavior. The study attempted to answer two ques-



tions: (a) If a student progresses through a series of school topics beginning with the least preferred and finishing with the most preferred, will his affinity for the least preferred topic be strengthened? (b) If a student progresses through a series of school topics beginning with the most preferred and finishing with the least preferred, will his affinity for the most preferred topic be weakened?

## 2. Subjects

One hundred and eight fifth-grade students at Cactus Wren Elementary School were randomly assigned to one of three groups. The average subject was 10 years of age. A majority of the subjects came from middle-income families. As far as *E* could determine, none of the subjects was mentally or physically abnormal. Slightly more than 50 percent were female.

## 3. Task

The experimental task required subjects to study three topics during a 90-minute daily session. The three topics were reading skills, science, and social studies. Following the completion of each assignment, the subjects were required to pass a test before progressing to the next topic. In order to pass the test, the subject had to achieve 100 percent correct responses. The nature of the tests varied among topics. In reading skills, the text was semi-programmed with objective questions. Correct answers were presented in the page margin so that students could check their own work. A teacher's approval was still required prior to progressing to the next topic. In social studies, essay questions were duplicated, and grading was done by the teachers. In science, experimental write-ups were graded by the teachers. Completed work was graded within five or 10 minutes.

## 4. Apparatus

This study required no formal apparatus. The materials used were the regular textbooks already in use at the school. In reading, *Reading Comprehension* (Sandford, Bishop, Gillespie, and Crosby) and *Vocabulary Development* (Deighton) were used. In social studies, *The Changing New World* (Silver Burdett) was used. The materials used in science were written by the teacher and involved descriptions of simple experiments to be completed by the students. A questionnaire was constructed by *E* to determine the subject's preferences for the three topics.

## 5. Procedure

The procedure can be separated into three distinct steps. First, all subjects were asked to complete a topic preference form. The second step of the study



was the treatment phase. The 90-minute treatment was in effect for 15 consecutive school days. The three groups each received different instructions. Prior to initiating the second phase, ratings obtained during the first phase were inspected to determine which was the least preferred (L), intermediate (M), and most preferred (H) topic for each subject in all groups. The third phase involved repeating the first phase as exactly as possible. All subjects were again asked to complete the questionnaire daily for five consecutive days. Exactly the same instructions were read again to all subjects.

Different instructions given to each treatment group during the second phase were as follows.

*a. Reward group.* All reward subjects were instructed to study the three topics in the given sequence each day. Each subject was then given a topic sequence such that he began with L, completed M, and finished with H. None of the subjects was told why he was to follow the given sequence.

*b. Punishment group.* All punishment subjects were instructed to study the three topics in the given sequence each day. Each subject was then given a topic sequence such that he began with H, completed M, and finished with L. No explanation was offered for the sequence.

*c. Control group.* All control subjects were instructed to study the three topics in whatever sequence they desired. These subjects were completely free of restrictions related to the sequence in which the topics were studied. Testing for these subjects was as lax as possible within the standards necessary for the student's normal academic progress.

### C. RESULTS

The preference data refer to the extent to which each subject "liked" each of the three topics: reading, social studies, and science. Scores are in the form of ratings on a 10-point scale. Subjects were instructed to use the following guideline to assign points to topics:

"I like it very much. It's my favorite" . . . Assign 8, 9, or 10.

"I like it a lot, but it's not my favorite" . . . Assign 5, 6, or 7.

"I like it" . . . Assign 2, 3, or 4.

"I neither like nor dislike it" . . . Assign 1.

"I don't like it" . . . Assign 0.

In general, the more a topic was liked, the more points it was assigned. The ratings were punched on cards and analyzed by a CDC 6400 computer.

To support the assumption that the ratings were a valid preference measure, subjects were also asked this question: "If you had one hour in which to study social studies, science, and reading, how many minutes would you

allot to each of these three topics?" In over 90 percent of the subjects, the ranking of topics on the basis of points was identical with the ranking of topics on the basis of time. Thus, if science received the most points, there was a strong tendency for science to be allotted the greatest amount of time.

In the first part of the analysis, three comparisons were made for each of the three groups. The reader is referred to Figure 1 for a display of the relevant comparisons. The symbols A-1 through A-6 will be used to identify the comparisons being presented.

Pretest

Posttest

## REWARD

L	. . . .Comparison	A-1. . . .	L
M	. . . .Comparison	A-2. . . .	M
H	. . . .Comparison	A-3. . . .	H

## PUNISHMENT

L	. . . .Comparison	A-4. . . .	L
M	. . . .Comparison	A-5. . . .	M
H	. . . .Comparison	A-6. . . .	H

## CONTROL

L	. . . .Comparison	A-7. . . .	L
M	. . . .Comparison	A-8. . . .	M
H	. . . .Comparison	A-9. . . .	H

FIGURE 1  
DIAGRAM OF COMPARISONS TESTED WITH SANDLER'S A STATISTIC

TABLE 1  
SUMMARY OF RATINGS

Topic	Pretest		Posttest		Mean difference
	Mean	SD	Mean	SD	
Reward group					
L (Least preferred)	3.2	2.65	4.8	3.26	+1.6
M (Intermediate)	6.9	2.50	6.8	2.88	— .1
H (Most preferred)	9.2	1.41	7.5	2.82	—1.7
Punishment group					
L (Least preferred)	2.1	1.98	3.8	3.02	+1.7
M (Intermediate)	6.6	2.15	6.5	2.76	— .1
H (Most preferred)	9.3	1.11	7.8	2.80	—1.5
Control group					
L (Least preferred)	1.9	2.33	3.6	3.16	+1.7
M (Intermediate)	6.0	3.27	6.6	3.51	+ .6
H (Most preferred)	9.4	1.58	8.1	2.75	—1.3

Mean pre- and posttest ratings for Reward subjects are given in Table 1. The major Reward group comparison was A-1. Sandler's *A* statistic was used to determine which of the pretest-posttest differences were statistically significant (6). Two comparisons were significant: A-1 ( $A = .11$ ;  $p < .01$ ), and A-3 ( $A = .08$ ;  $p < .01$ ).

Mean pre- and posttest ratings for Punishment group subjects are given in the second part of Table 1. The most important comparison was A-6. Two comparisons were statistically significant: A-4 ( $A = .08$ ;  $p < .01$ ), and A-6 ( $A = .12$ ;  $p < .01$ ).

Mean pre- and posttest ratings for Control subjects are also given in Table 1. Two comparisons were significant: A-7 ( $A = .11$ ;  $p < .01$ ), and A-9 ( $A = .12$ ;  $p < .01$ ).

The two significant comparisons of major importance involved the Reward group's least preferred topic and the Punishment group's most preferred topic. Figure 2 presents a comparison between the Reward group's least preferred topic and the corresponding Control group's least preferred topic. Figure 3 presents a similar comparison involving the Punishment group's most preferred topic and the corresponding Control group's most preferred topic.

To determine whether the significant comparisons noted above (A-1, A-3, A-4, and A-6) could actually be related to the contingency, four *t* tests were completed. The reader is referred to Figure 4 for a display of the relevant comparisons. None of the *t* tests was significant.

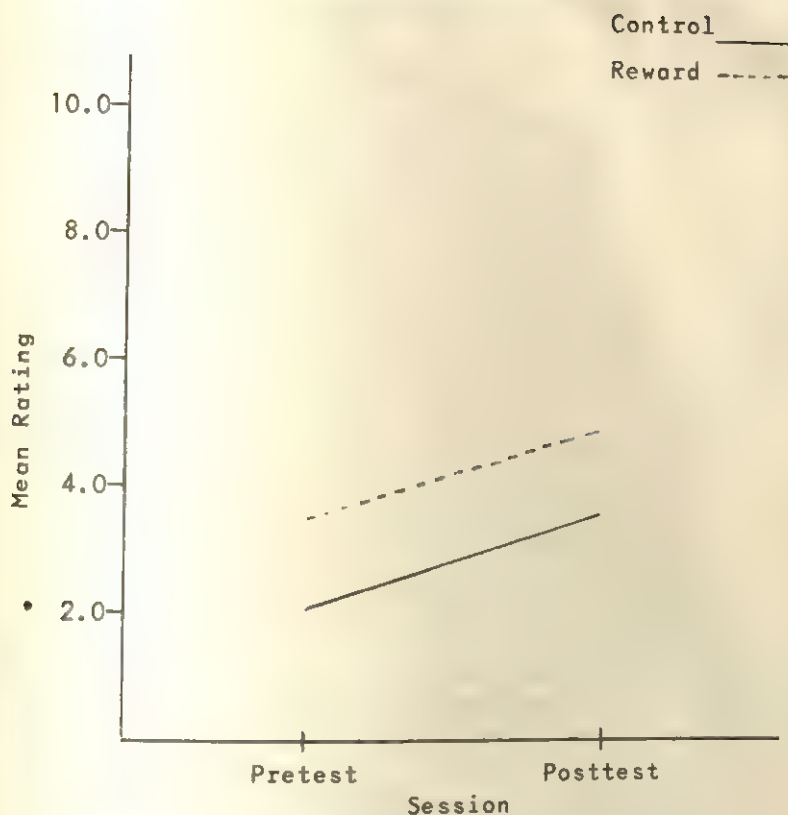


FIGURE 2  
MEAN CHANGES IN TOPIC PREFERENCES FOR LEAST PREFERRED TOPIC

#### D. DISCUSSION

In the literature surveyed prior to this experiment, it was noted that tests of the Premack principle had usually involved animal subjects. The most notable exception to this tendency relevant to the present experiment involved several studies completed by Wasik (8, 9, 10). Although most of her results supported the Premack principle, there was one negative finding. In a test involving multiple contingencies, Wasik (8) found that it was possible to reinforce a given response by following that response with the opportunity to perform a *less* probable behavior. This paradox led Wasik to conclude that the principle was valid only for response pairs and not

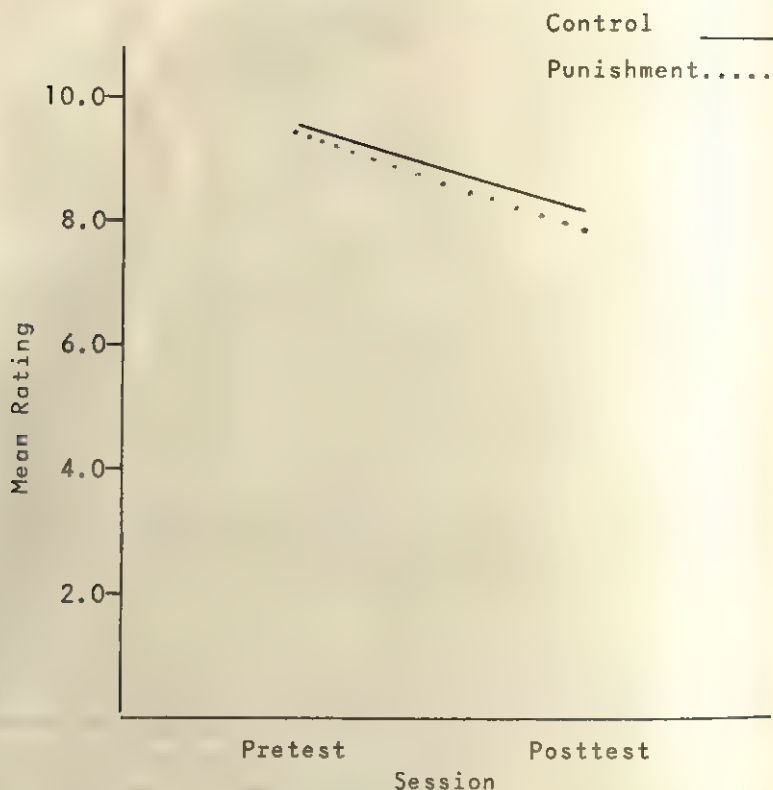


FIGURE 3  
MEAN CHANGES IN TOPIC PREFERENCES FOR MOST PREFERRED TOPIC

valid for response chains involving multiple contingencies. This conclusion contradicted the results of a study by Premack (3) in which monkeys were presented a choice among four manipulanda. On the basis of choices, a response preference hierarchy was established (A, B, C, D). The most preferred response (A) was found capable of reinforcing all others; the least preferred (D) reinforced no others. Response B reinforced C and D, response C reinforced only D. Thus, all of the reinforcing effects conformed to the theoretical predictions.

The present study tested the reinforcing capabilities of a Premackian multiple contingency in a classroom setting. The results indicate that the contingency had no effect upon topic preferences. One of the several possible conclusions is that Wasik's interpretation was essentially valid; e.g., the



Reward Posttest		Control Posttest
L	. . . . . Comparison T-1. . . . .	L
H	. . . . . Comparison T-2. . . . .	H
Punishment Posttest		Control Posttest
L	. . . . . Comparison T-3. . . . .	L
H	. . . . . Comparison T-4. . . . .	H

FIGURE 4  
COMPARISONS TESTED WITH STUDENT'S *t* TEST

Premack principle is simply not applicable in its present form to multiple contingencies. Further research in an environment subject to greater control will be necessary to determine the applicability of the principle to multiple contingencies.

An alternative explanation for the failure of the contingency to have a reinforcing effect involves the setting of the present study. The research spanned a five-week period. During the treatment phase, Control subjects were supposedly free to study their topics in whatever sequence they desired. Ideally, Control subjects could choose to devote a whole 90-minute session to their favorite topic. Unfortunately, this was not the case. Because of the practical necessities associated with acquiring an education, the teachers quite understandably insisted that Control subjects spend some time each day on all three topics. This requirement tended to nullify the distinction between the two experimental groups and the Control group. In addition, Control subjects tended to work in pairs with Reward and Punishment subjects. In order to work together, both subjects had to study the topics in the same sequence. This resulted in the Control subjects tending to follow the same sequence as their partner, which also served to cloud the experimental-control group distinction.

To the present author, the results indicate that statistical regression alone could account for the changes. In each of the three groups, only the extreme scores (L and H) showed significant change. Also, in all cases the direction of change was towards the mean; low scores increased, high scores decreased. This is exactly the type of change one would predict to result from selecting scores on the basis of extremeness.

The goals of the present study were twofold: (a) to test the theoretical validity of the Premack principle with multiple contingencies, and (b) to determine whether the principle could be applied to a classroom situation involving low probability behaviors. Neither goal was clearly realized, but negative results do not disprove a theory. In order to realize the first goal, a test will have to be made under much more closely controlled conditions. The reinforcing effect of the contingency is in part dependent upon the strength of the contingent behavior. If the strength is low, the reinforcing effect will be relatively weak. A high degree of control over random variation will be prerequisite to demonstrating such an effect. In the present study, over 100 fifth graders and four teachers combined with the necessity of conforming to standard classroom procedures to generate a high level of random variance which probably obscured any potential treatment effects.

The fulfillment of the second goal will require a more powerful reinforcer. Fifth-grade students are simply not reinforced by being allowed to study certain topics. If the contingent behavior was recess, a stronger effect would undoubtedly be evident. Thus, in order to test the principle with low strength behaviors, one must employ a highly controlled environment. To modify classroom topic preferences, one must use a more powerful reinforcer.

Premack (5) also attempted to use his reinforcement model to explain the basis of punishment. If reinforcement is basically any situation in which higher strength behavior is contingent upon lower strength behavior, punishment may simply be the reverse; lower strength behavior contingent upon higher strength behavior. Since the individual will not freely initiate a less probable behavior, punishment requires some kind of force to insure that the contingency is realized.

In the present study, Punishment subjects began each day's study with their most preferred topic. At the end of 30 minutes, they were required to move on to their intermediate topic. It was predicted that this procedure would result in a decrement in the strength of the most preferred topic. Although a significant decrement was realized, a similar decrement also was evident for Control subjects. For this reason, the Punishment decrement cannot be attributed to the reverse contingency. However, it would be

erroneous to conclude that Premack's notion of punishment is invalid on the basis of these results. It is more probable that the same conditions which were cited to explain the failure of the reinforcing contingency are applicable in the case of punishment. The punishing effect is relatively weak and will only be detected when extraneous and overpowering influences are reduced.

The Premack principle purports to be applicable to behaviors ranging in probability from low to high. When the principle is being tested with relatively low strength behaviors, it predicts that the reinforcing effect will be weak. In order for the reinforcing potential actually to become evident, extraneous sources of interference must be kept minimal. This suggests that the principle will be found to have little practical effect upon classroom behaviors when the contingent responses are low strength. The present procedure could probably produce practical results if higher strength responses were in effect. If, on the other hand, the research is only concerned with validating the principle, a highly controlled environment is suggested.

For further information on the relevance of the principle to classroom behavior, this experiment should be systematically replicated with use of progressively stronger and stronger contingent responses. This would establish a "probability threshold," marking the minimum contingent response strength prerequisite to effective classroom reinforcement.

It would also be of interest to replicate the study with use of the same treatment but a different setting which allowed more control. One group of college students, for example, could be asked to study their assignments in specified sequences, while control subjects were completely free of any such restrictions. This would, in part, overcome the limitations cited in the present study.

#### REFERENCES

1. DOLMATCH, T. B. *Revolution in training: Programmed instruction in industry*. New York: Amer. Manag. Assoc., 1962.
2. GLASER, R. *Training Research and Education*. Pittsburgh, Pa.: Univ. Pittsburgh Press, 1961.
3. PREMACK, D. Reinforcement theory. In D. Levine (Ed.), *Nebraska Symposium on Motivation*. Lincoln: Univ. Nebraska Press, 1965. Pp. 123-180.
4. ———. Catching up with common sense or two sides of a generalization: Reinforcement and punishment. Paper read at a Conference on the Nature of Reinforcement, Center for Research in Human Learning, University of Pittsburgh, June, 1969.
5. PREMACK, D., & SHANAB, M. Rats prefer the home cage to the runway following intermittent but not consistent reinforcement. *Nature*, 1968, **217**, 288-289.
6. RUNYON, R. P., & HABER, A. *Fundamentals of Behavioral Statistics*. Reading, Mass.: Addison-Wesley, 1967. P. 172.

7. SKINNER, B. F. Teaching machines. *Science*, 1958, **128**, 969-977.
8. WASIK, B. H. Tests of the applicability of Premack's generalization of reinforcement to human lever-pressing behavior. Unpublished Doctoral dissertation, Florida State University, Tallahassee, 1967.
9. ———. A postcontingency test of the effectiveness of reinforcement. *Psychon. Sci.*, 1968, **13**, 87-88.
10. ———. The application of Premack's generalization on reinforcement to the management of classroom behavior. *J. Exper. Child Psychol.*, 1970, **10**, 33-43.

*Graduate School of Social  
Service Administration  
Arizona State University  
Tempe, Arizona 85281*

## CONSERVATION OF SPACE IN NONINSTITUTIONALIZED OLD PEOPLE\*

*The University of Wisconsin*

---

DIANE E. PAPALIA, ELIZABETH KENNEDY, AND NANCY SHEEHAN

---

### SUMMARY

Forty-eight noninstitutionalized elderly were administered Piaget's Conservation of Surfaces Test. Only 13 subjects were designated as conservers in this content domain. A strong relationship existed between conservation ability and educational attainment, but not between sex and performance, and age and performance. Results indicated that cognitive abilities may be subject to qualitative disorganization in the latter portions of the life-span.

---

### A. INTRODUCTION

The extent of conservation ability during late adulthood has not been determined conclusively. Furthermore, previous investigators (e.g., 4, 10) who studied this concept in the elderly failed to control for any possible effects that living condition (i.e., institutionalization *vs.* noninstitutionalization) might have on performance. Sanders *et al.* (10) based their conclusion that there is deterioration of the concept of space with advanced age on performance of noninstitutionalized and institutionalized elderly. No attempt was made to disentangle the effects of living arrangement from performance. Kominski and Coppinger (4) made similar conclusions on the basis of samples of young and old adult males living in a Veterans Hospital. However, Weinstock and Bennett (13) noted that prolonged institutional residency tends to depress cognitive proficiency.

The present report is of an investigation of space conservation performance in a group of active, healthy old people who resided in private homes or apartments and were regular participants in senior citizen club activities. None was under any type of custodial care.

---

\* Received in the Editorial Office on February 28, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.



## B. METHOD AND PROCEDURE

1. *Subjects*

Ss were 48 Caucasian middle-class adults (23 females and 25 males). Ages of the subjects ranged from 63 to 92 years; mean age was 74.4 years ( $SD = 6.6$ ). Educational attainment ranged from 5 to 19 years; mean educational level was 10.7 years ( $SD = 3.2$ ).

Ss were current participants in the activities of senior citizens' centers in and near Madison, Wisconsin. All Ss were contacted and tested individually at the center. Ss were free from marked auditory and visual impairment.

2. *Procedure and Test Materials*

One female experimenter individually administered to all Ss Piaget's Conservation of Surfaces Test to evaluate the concept of space in the elderly. Ss were given three conservation of space trials and a trial-and-check in which an inequality was to be judged. Ss were given as much time as necessary to complete the task.

In the test, two pieces of green cardboard ( $9 \times 12$  inches) were presented to the subject, and *E* explained as follows [adapted from Shantz and Sigel (11)]:

Let's pretend that these boards are two fields of grass. They are the same size, see. If we put a cow in each field, each cow has just as much grass to eat as the other cow [*E* places a barn in the left field]. Now Farmer Jones builds a barn on this field. He has to take some of the grass away to make room for the barn. Now show me which cow has more grass to eat. Yes (or no), that [*E* pointing] cow has more grass to eat, because the barn covers up part of this cow's grass. Now I will put a barn on the other field. Now do both cows have the same amount to eat? [*E* puts both barns in upper left-hand corner].

Three, nine, and six barns were used in trials 1 through 3, respectively. In the trial-and-check, five barns were placed on the left board and four on the right. Each trial was preceded by these instructions: "Watch what I do. You see, I am putting some more barns in each field. Now does this cow have as much grass to eat as that one, or does one have more grass to eat?" Following each trial, Ss were required to explain their response. Ss were designated as conservers under two scoring conditions. In the first condition ("justification condition"), performance was assessed on the basis of Ss' objective response and its adequate justification. Adequate justifications were those that indicated that a logical thought sequence was used in arriving at a response [see Papalia and Hooper (7)]. In the second condition ("no-

justification condition"), Ss were judged on their objective response alone and were not required to explain their reasoning. In both conditions, Ss were classified as "conservers" if they passed at least three trials; all other Ss were considered to be "nonconservers." An overall performance score (range: 0-4) was also assigned to each S.

### C. RESULTS AND DISCUSSION

Initial analysis indicated no difference in number of Ss passing on the basis of scoring criteria employed ( $\chi^2 = 2.25$ ;  $p > .20$ ). Consequently, all further analyses were based upon the "justification condition" data. Since no significant sex difference was found ( $\chi^2 = .02$ ;  $p > .99$ ), combined male-female data were used for subsequent analyses.

Of the total sample ( $N = 48$ ), 13 Ss (six males, seven females) conserved. The remaining 35 Ss (17 males, 18 females) did not attain a score of 3 or 4 on the task.

Pearson product-moment correlation coefficients were computed between performance (score: 0-4) and age, and between performance and years of education. A significant inverse relationship between age and performance existed ( $r = -.36$ ;  $p < .02$ ), and a significant positive relationship was found between educational level and performance ( $r = .35$ ;  $p < .02$ ). The correlation between age and education was significant and positive ( $r = .74$ ;  $p < .001$ ). A partial correlation between age and performance, with years of education held constant, significantly altered the apparent negative correlation between age and performance ( $r = .17$ ; n.s.). However, a partial correlation between performance and education, with age held constant, yielded a markedly higher correlation between these two variables ( $r = .98$ ;  $p < .001$ ). Thus, educational level appeared to have a crucial relationship with cognitive status in the elderly.

These results indicate that the basic trends noted by Sanders *et al.* (10) and Kominski and Coppinger (4) are characteristic of active, healthy, non-institutionalized old people. With acknowledgment of the limitations of cross-sectional data, tentative support is provided for the hypothesis of Hooper, Fitzgerald, and Papalia (3) that logical thought may be subject to regression and qualitative disorganization with age.

According to certain theoretical speculations (e.g., 2, 9), the abilities tapped by "Piagetian" measures, once they have entered a person's response repertoire, should be maintained throughout the remainder of the life-span. However, recent empirical examinations suggest that this may not be the case. Less than optimal cognitive performance in the elderly has been noted

in the domains of conservation ability (4, 6, 8, 10, 12), egocentrism (5), and classification skills (1, 12).

Attention must now be directed toward assessing which variables contribute to the maintenance or decline of cognitive abilities in later life. One particularly salient variable appears to be level of formal schooling. However, the relationship between educational attainment and cognitive performance in adulthood has been inconsistent. Several recent studies have noted a nonsignificant relationship between these variables (e.g., 6, 10, 13). On the other hand, in Papalia, True, and Salverson (8), as well as in the present research, a positive, generally significant relationship between conservation performance and education was reported. Significant differences in amount of education existed among the various samples ranging from 5.3 years (10) to approximately 11 years (6, 8, 12, and the present sample); such differences may bear on the present controversy. Further research should focus on the characteristics of highly educated *vs.* poorly educated individuals (e.g., relationship between amount of education and initial ability level, present activity level, socioeconomic status and related occupational level, degree of anxiety in testing situations, and certain personality dimensions) to uncover those underlying variables that contribute to the status of cognitive abilities in the elderly.

#### REFERENCES

1. DENNEY, N., & LENNON, M. Classification: A comparison of middle and old age. *Devel. Psychol.*, 1972, 7, 210-213.
2. FLAVELL, J. Cognitive change in adulthood. In L. R. Goulet & P. B. Baltes (Eds.), *Life-Span Developmental Psychology: Theory and Research*. New York: Academic Press, 1970.
3. HOOPER, F., FITZGERALD, J., & PAPALIA, D. Piagetian theory and the aging process: Extensions and speculations. *Aging & Hum. Devel.*, 1971, 2, 3-20.
4. KOMINSKI, C., & COPPINGER, N. Piaget's conservation of space and the Muller-Lyer illusion in the elderly. Unpublished manuscript, College of William and Mary, Williamsburg, Virginia, 1968.
5. LOOFT, W., & CHARLES, D. Egocentrism and social interaction in young and old adults. *Aging & Hum. Devel.*, 1971, 2, 21-28.
6. PAPALIA, D. The status of several conservation abilities across the life-span. *Hum. Devel.*, 1972, 15, 229-243.
7. PAPALIA, D., & HOOPER, F. A developmental comparison of identity and equivalence conservation. *J. Exper. Child Psychol.*, 1971, 12, 347-361.
8. PAPALIA, D., TRUE, M., & SALVERSON, S. An evaluation of mass, weight, and volume conservation ability during old age. Paper presented at the annual meeting of the Gerontological Society, San Juan, Puerto Rico, December 1972.
9. PIAGET, J., & INHELDER, B. *The Psychology of the Child*. New York: Basic Books, 1969.

10. SANDERS, S., LAURENDEAU, M., & BERGERON, J. Aging and the concept of space: The conservation of surfaces. *J. Gerontol.*, 1966, 21, 281-285.
11. SHANTZ, C., & SIGEL, I. Logical operations and concepts of conservation in children: A training study. Center for Developmental Studies in Cognition, The Merrill-Palmer Institute, Detroit, Michigan, 1967.
12. STORCK, P., LOOFT, W., & HOOPER, F. Interrelationships among Piagetian tasks and traditional measure of cognitive abilities in mature and aged adults. *J. Gerontol.*, 1972, 27, 461-465.
13. WEINSTOCK, C., & BENNETT, R. From "waiting on the list" to becoming a "new-comer" and an "old timer" in a home for the aged: Two studies of socialization and its impact on cognitive functioning. *Aging & Hum. Devel.*, 1971, 2, 46-58.

*Department of Home Management and Family Living*  
*University of Wisconsin-Madison*  
*1270 Linden Drive*  
*Madison, Wisconsin 53706*





## EFFECTS OF EXTERNAL GENITAL SENSORY FEEDBACK ON COPULATORY BEHAVIOR OF RATS\*

*Kennedy-King College*

---

PETER O. PERETTI

---

### SUMMARY

The role of external genital sensory feedback and sex of *S* and the effects on copulatory behavior were investigated. The prevention of semen plugs and anesthetization of the vagina had no significant effect on the running times of the female or the number of mounts of the male, when compared to a control group. Anesthetization of the penis and prevention of erection significantly reduced the number of mounts made by the male. Short term changes in external genital sensory feedback were found to be one factor affecting copulation; however, nonsexual factors were also suggested as possible determinants and cited as areas of subsequent research.

---

### A. INTRODUCTION

The normal copulatory pattern of rats seems to consist of a series of brief penetrating contacts (intromissions) separated by intercopulatory intervals varying from 20 seconds to 2 minutes in different strains (10). After eight to 15 intromissions the male ejaculates. Following the initial ejaculation, the male delays for approximately 5 minutes before making another intromission. The number of intromissions in the second ejaculatory series and all subsequent series is always fewer than that in the first series (3, 5, 8).

Bermant and Westbrook (6) found that there is a clear inverse relationship between the intensity of a copulatory stimulus and the rate at which females will produce further stimulation. They also found it possible to decrease the intromission intervals and thus increase the rate of intromission to the point where the vagina bled.

Pierce and Nuttal (9) have shown that sexual contact can be aversive for the female rat. A female rat in estrous was placed in a group of severely sexually aggressive males and was found to perform an escape response (pressing a lever) to escape subsequent copulatory contact. The probability

---

\* Received in the Editorial Office on March 1, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

and duration of the responses was least after mounts and greatest after ejaculation.

Internal sensory feedback is defined as the stimulus input arising directly from the females' own responses. Several experimentors have shown that internal sensory feedback does not affect the basic receptivity. Beach (2) has shown that female rats, in which the ovaries, tubes, uterus, and vagina were congenitally absent, could execute the copulatory pattern under the influence of exogenous hormones. Ball (1) showed that fundamental sexual receptivity is not dependent for its maintenance upon genital stimulation.

Bermant (4) has shown that estrous female rats will press a lever if this is followed by the presentation of a male rat, and that the females will establish through their lever pressing a set of "preferred intervals" for copulation. Pierce and Nuttal (9) also examined the contact-response interval variable with the same results. These studies indicated that female Ss develop a pattern of responding in which the interval of time from a single sexual contact to the next contact-producing response depends strongly on the type of contact.

The scope of the forthcoming experiment was to examine a number of variables which may affect rats in their copulatory behavior patterns. Of primary interest was the role of external genital sensory feedback as a factor affecting the male-female contacts.

## B. METHOD

### 1. Subjects

Ten Sprague Dawley female Ss approximately 120 days old were obtained and maintained on schedules for water and food. Ten male Ss of the same strain were also used and maintained on the above schedule. Female and male Ss were matched in age and strain.

### 2. Apparatus

A straight maze  $24\frac{1}{2}'' \times 3\frac{1}{2}'' \times 3\frac{1}{2}''$  was used. The starting box was  $4'' \times 3\frac{1}{2}'' \times 3\frac{1}{2}''$ . The box had two slide openings; the rear one allowed the S to be placed into the box, and the forward opening (the starting gate) opened to the runway. Slides were  $3'' \times 5''$  and made of cardboard.

The runway followed directly into a cubicle at the end (the goal box)  $10'' \times 8'' \times 11''$ . A barrier between the runway and the cubicle was made of glass which allowed the S to see beyond it and which could be removed.

### 3. Procedure

For a period of two weeks prior to each testing period each female *S* was housed with one male. It was assumed that since the female *Ss* would have been estrous at least once during this period, both the males and females would no longer be sexually naive. After this time period, the male and female *Ss* were divided into two groups of five each (one experimental group and one control).

Female experimental *Ss* were brought into estrous by combination injections of .1 mg of estrogen (Upjohn, Depo-Estradiol), intramuscularly, 24 to 30 hours prior to testing and .5 mg progesterone (Upjohn, Progesterone), intramuscularly, 4 to 6 hours prior to testing. In all subsequent testing, the same procedure was used to bring about estrous. The interval between injections for each treatment was maintained at two weeks. This minimized carry-over from one treatment to the next. *Ss* were determined to be in estrous if they had reached stage 10 of the Hemmingsen test (7) for sexual arousal (ear wiggling and lordosis in response to digital stimulation of the genital region).

There were five different treatments used in this experiment. The five estrous females were used in all treatments, but in every different treatment were treated differentially. Therefore no continuity or order-effect was assumed in the treatment of the results, thus making the equivalent of a total of 25 experimental *Ss* used. For each treatment each of the five *Ss* made eight trial runs, and the running time was recorded as the time from the opening of the goal box to the time the *Ss* had completely crossed into the goal box. The number of mounts the males made were also recorded. A mount was defined as any successful attempt on the part of the male to grasp the female's flanks or back and hold his head and shoulders over the female's back. Control *Ss* were run in the same way. They received *no* chemical treatments.

a. *Treatment 1.* The first treatment consisted of placing the male into the goal box prior to the beginning of testing. Three different goal boxes were used at the same time, each containing an incentive male. The runway was then slid up to the first goal box, and the estrous female was placed in the start box. Once the female was approximately 4" from the goal box and did not appear to hesitate, the glass partition was opened. Once the female *S* was inside the goal box, the glass partition was replaced, and the runway was slid away and to the second goal box. The *Ss* remained in the first box for a

total of 10 minutes, and the number of mounts made by the male was recorded. While this was occurring, a second female was placed into the start box, and the same procedure carried through with another male *S*. Experimental animals and the controls were run intermittently. The apparatus was lightly cleaned with soap and water after each group of eight runs to help avoid olfactory cues.

*b. Treatment 2.* Treatment 2 consisted of giving the male incentive 50 mg/kg intraperitoneal injections of guanethedine sulfate (Ciba, Ismelin) 3 to 4 hours prior to testing. Intraperitoneal injections at this dosage level prevents the expulsion of semen during the ejaculatory reflex without otherwise significantly altering the qualitative aspects of male rats' sexual behavior (6). The same procedure described in treatment 1 was then carried through. All estrous *Ss* in this treatment were matched with the same male incentive they went towards in treatment 1.

*c. Treatment 3.* Treatment 3 consisted of swabbing the male incentive's genital area, particularly the glans penis and shat with 4% xylocaine hydrochloride paste (Astra, Xylocaine), approximately 3 to 5 cc, 20 minutes prior to testing. The genital area was then swabbed with cotton, and following a 1 minute delay, a similar application was made. At a 4% concentration, xylocaine paste is a rapid acting and long sustained topical anesthetic. It was used in this treatment to prevent the erection of the penis. Following the testing in treatments 3, 4, and 5, the genital area was swabbed with cotton and washed with water. The same procedure of testing as in the previous treatments was carried through.

*d. Treatment 4.* Treatment 4 required the same type of application of xylocaine paste to the genital area of the female. The exact depth of anesthesia was at least the first 4 mm of the vagina. The same procedure as described above was then used.

*e. Treatment 5.* Treatment 5 consisted of swabbing the vaginal area according to the method described in treatment 4 with vaseline. Vaseline was chosen, since it is similar in consistency to the xylocaine paste, yet it has little affect on the vagina when applied. *Ss* were run as above.

### C. RESULTS

Running time was the main variable recorded and was counted as the time the starting gate opened until the *S* had crossed into the goal box. Eight trials were run with each *S* in all five treatments. Table 1 is a summary of the total time of each female experimental *S* in each treatment and also the

TABLE 1  
TOTAL AND MEAN NUMBER OF SECONDS OF RUNNING TIME OVER EIGHT TRIALS  
FOR THE FIVE FEMALE EXPERIMENTAL Ss IN EACH OF THE FIVE TREATMENTS

Treatment	Running time by subject									
	S-1		S-2		S-3		S-4		S-5	
	Total	Mean	Total	Mean	Total	Mean	Total	Mean	Total	Mean
1	49.6	6.2	31.7	3.96	57.9	7.24	37.7	4.71	48.4	6.05
2	41.2	5.15	51.1	6.26	44.8	5.60	45.9	5.74	41.7	5.21
3	42.7	5.34	54.2	6.78	52.6	6.33	58.5	7.31	40.7	5.88
4	46.0	5.75	74.5	9.31	53.2	6.65	43.8	5.48	52.4	6.55
5	60.5	7.56	45.0	5.68	46.4	5.80	64.4	8.05	69.6	8.70

mean running time for the eight trials. The overall mean running time was 6.26 seconds.

An analysis of variance was used to determine the differences for experimental Ss between the varying treatments. The  $F$  of .2921 (4 and 20  $df$ ) was not significant. Multiple  $t$  tests were also run to compare the effect of the different treatments between the experimental and control Ss. The results are in Table 2. Minimum significance levels are in the lower half of the same table.

The number of mounts made during the 10 minute postrunning period were also recorded. A total of 484 mounts were made by the 25 pairs of experimental Ss. Table 3 summarizes the total number of mounts made during the 10 minute period during the eight trial runs of each experimental male  $S$  and the mean number of these mounts per trial. An analysis of variance between treatments showed that  $F$  was equal to 12.0377 ( $df = 4, 20$ ), and the significance was beyond the .01 level. Scheffé's multiple comparison *post hoc* analysis was then used to compare the treatment means. The alpha level was set at .05, and the mean difference needed for significance was found to be .98.

Table 4 is a summary of these means with significant differences in italics.

TABLE 2  
 $t$  TEST COMPARISONS MADE BETWEEN DIFFERENT TREATMENTS AND THE  
FEMALE EXPERIMENTAL vs. FEMALE CONTROL Ss, AND SIGNIFICANCE LEVELS  
OF EACH COMPARISON (ON LOWER HALF)

Subject	Treatment				
	1	2	3	4	5
S-1	—	.037	.767	1.227	1.751
S-2	.32	—	1.038	1.530	2.210
S-3	.25	.20	—	.632	1.197
S-4	.15	.10	.30	—	.461
S-5	.10	.05	.15	.35	—



TABLE 3  
TOTAL NUMBER OF MOUNTS MADE BY EACH MALE EXPERIMENTAL S DURING  
EIGHT TRIALS IN EACH TREATMENT AND THE MEAN NUMBER OF  
MOUNTS PER 10 MINUTES POSTRUNNING TIME PERIOD

Subject	1		2		Treatment 3		4		5	
	Total	Mean	Total	Mean	Total	Mean	Total	Mean	Total	Mean
S-1	18	3.6	24	4.8	11	2.2	26	5.2	20	4.0
S-2	23	4.6	17	3.4	19	3.8	21	4.2	19	3.8
S-3	21	4.2	28	5.6	14	2.8	16	3.2	18	3.6
S-4	19	3.8	21	4.2	7	1.4	19	3.8	22	4.4
S-5	16	3.2	21	4.2	20	4.0	24	4.8	20	4.0

The means used were obtained by summing the treatment means in Table 3 and are shown in Table 4. Four comparisons were found to be significant at the .05 level: 3-1, 3-2, 4-3, and 5-3.

#### D. CONCLUSION AND DISCUSSION

The results of this experiment indicated that the reduction of sensory stimulation had no significant effect on the running times of the Ss, but did significantly reduce the number of mounts made. Running times of the Ss in general did not significantly increase or decrease under the varying conditions. However, in one comparison of the treatments—i.e., between treatment 2 (guanethedine sulfate) and treatment 5 (vaseline)—running time did show a significant (.05) difference.

It was assumed in view of the work of Bermant and Westbrook (6) and Pierce and Nuttal (9) that there is an inverse relationship between intensity of stimulation and desire for further stimulation. Thus the Ss in treatments 2, 3, and 4 should have traveled the runway faster than those Ss in treat-

TABLE 4  
MEAN NUMBER OF MOUNTS PER 10 MINUTE PERIOD MADE BY MALE EXPERIMENTAL  
SS IN EACH TREATMENT AND THE MEAN DIFFERENCES FOUND

Subject	Treatment				
	1	2	3	4	5
All	3.88	4.44	2.84	4.24	3.96
		<i>Mean mounts</i>			
		<i>Mean differences</i>			
S-1					
S-2	.68				
S-3	1.40	1.60			
S-4	.36	.20	1.40		
S-5	.08	.48	1.12	.28	

Note: Statistically significant differences shown in italics.

ments 1 and 5. However, in view of the results this assumption is questioned; the reductions of stimulation did not increase the times.

The number of mounts made during the 10 minute postrunning period did not significantly increase for any of the treatments. It was assumed that an increase in copulatory activity would follow a decrease in sensory feedback. This assumption was rejected, and the opposite was found true, especially for the treatment in which the penis was anesthetized. For any generalization to be made from these results, it might be considered that the number of mounts made by the male is not simply a reflection of the copulatory reflexes but more of a reflection of the increasing levels of excitement in both the male and female.

It was concluded that reduction of nongenital stimulation in this experiment had little effect on the need for contact in the pairs, contrary to what might have been expected. The conclusions drawn by Bermant and Westbrook (6) and Pierce and Nuttal (9) have not been disproven. Their experiments allowed the female a greater control of the time factors and male presentation which could affect results.

It may also be concluded that only the prevention of the male's erection had an effect on the quantitative and qualitative aspects of the copulatory contact. The prevention of the plug formed in the vagina from the semen and the anesthetization of the vagina had no significant effect on the number of mounts made. General conclusions drawn in this area suggest that the prevention of erection does not cause a qualitative reduction in copulatory contact. For this reason, it would seem likely that an unknown variable affected the males' performance in this treatment.

Even though the female Ss used in this experiment were matched in age and strain, a large amount of variation in behavior was already present. It could be that these existing variations were more important than the variables measured. For example, it might be reasoned that if copulatory behavior had an appetitive nature for the female rats, then the artificial induction of estrous could quite possibly have had an effect on the females' performance. It would appear that measurement of what motivates copulatory behavioral patterns would be more important after the basic physiological and behavioral patterns were known, rather than measuring the results of physiological changes.

Because of the limited knowledge in the social, affective, and communication variable, any attempt at changing a physiological variable could possibly result in a change of a great number of other variables which are difficult to measure. Thus, subsequent research concerned with sexual motivation

might deal primarily with the interaction of a large number of nonsexual stimuli and with what nonphysiological stimuli motivate copulatory behavior.

## REFERENCES

1. BALL, J. Normal sex behavior in the rat after total extirpation of the vas deferentia. *Anat. Rec.*, 1934, **58**, 49.
2. BEACH, F. Copulatory behavior in prepuberally castrated male rats and its modification by estrogen administration. *Endocrinology*, 1942, **31**, 679-683.
3. BEACH, F., & JORDAN, L. Sexual exhaustion and recovery in the male rat. *Quart. J. Exper. Psychol.*, 1956, **8**(3), 87-123.
4. BERMANT, G. Response latencies of female rats during sexual intercourse. *Science*, 1961, **133**, 1771-1773.
5. ———. Effects of single and multiple intercopulatory intervals on the sexual behavior of male rats. *J. Comp. & Physiol. Psychol.*, 1964, **57**, 498-503.
6. BERMANT, G., & WESTBROOK, W. Peripheral factors in the regulation of sexual contact by female rats. *J. Comp. & Physiol. Psychol.*, 1966, **61**, 244-250.
7. HEMMINGSEN, A. Studies in oestrus-producing hormone (oestrus). *Scand. Arch. Physiol.*, 1933, **65**, 97-250.
8. LARSSON, K. Conditioning and Sexual Behavior in the Male Albino Rat. Stockholm: Almquist & Witsell, 1956.
9. PIERCE, J., & NUTTAL, R. Duration of sexual contacts in the rat. *J. Comp. & Physiol. Psychol.*, 1961, **54**, 585-587.
10. WHALEN, R. Strain differences in sexual behavior of the male rat. *Behavior*, 1961, **18**, 199-204.

Department of Psychology  
Kennedy-King College  
6800 South Wentworth Avenue  
Chicago, Illinois 60621

## PRONUNCIABILITY IN VERBAL LEARNING OF THE DEAF\*<sup>1</sup>

*Rochester Institute of Technology*

---

KATHLEEN CHEN

---

### SUMMARY

The investigation was concerned with the effect of rated pronunciability on learning of trigrams by three categories of deaf subjects differing in histories of hearing loss. Lists of trigrams of three levels of difficulty were presented through a memory drum to the three groups of 18 Ss each. A total of four free recall trials was given for learning of each list.

By means of the Friedman two-way analysis of variance and Mann Whitney *U* test, the following were the results: (a) Pronunciability is a predictor of learning for all categories of deafness. (b) With a moderate level of difficulty, the profoundly deaf Ss performed significantly more poorly than those in other categories.

---

### A. INTRODUCTION

Study by Conrad and Rush (2) on memory encoding of the deaf confirmed the absence of acoustic encoding of letters in deaf subjects, with the mode of deaf encoding processes still unclarified. Conrad (1) provided evidence of different coding types on the part of the deaf children as compared to hearing subjects. Deaf children learned a list of homophone word pairs and a list of pairs of words of similar shape equally well. His investigation provided evidence that deaf children have different coding types. The main purpose of the present study was to extend our knowledge concerning the variables that affect the verbal learning of the deaf. Underwood and Schulz (5) and Underwood and Postman (4) had found that ease of paired-associate and serial learning was closely related to pronunciability ratings (PRs) with hearing subjects. The only negative case was that of Lindley (3). Do the deaf subjects also learn a trigram by telescoping it into a simplified

---

\* Received in the Editorial Office on March 2, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was supported by a National Science Foundation Institutional Grant. The author would like to express her appreciation to Dr. Gerard G. Walter of the National Institute for the Deaf at Rochester Institute of Technology, for providing the demographic data of the deaf students and for his very capable assistance in selecting the deaf students.

sound unit, or do they learn it as a series of three independent letters? If the former is the case, then rated pronunciability, which may reflect the ease of articulating the unit as a whole, may be a predictor of verbal learning of the deaf. The present experiment investigated the role of pronunciability on the learning of the different categories of deaf subjects with histories of varying hearing loss.

Another consideration is the differential rate of learning of the different categories of deaf subjects. Differences in performance might be explained by an auditory frequency hypothesis. One dimension of a verbal unit is frequency, defined as the number of times that a given verbal item is found in printed material or as a part of the subject's oral repertoire. It has been assumed that objective frequency reflects the experience that a subject has had with the material. The present study used subjects with different histories of deafness. Hence the visual frequency of experience of all groups is constant for the three groups, except for the variable of auditory frequency of experience, both received and emitted, which differs in terms of the extent of hearing loss and age of onset of deafness. It was reasoned that rate of hearing of subjects with less hearing loss will be greater than those with profound hearing loss.

## B. METHOD

### 1. Lists

Three lists of trigrams varying in PRs were used. The estimate of pronunciability was the Underwood and Schulz (5, Appendix E) PRs. In these ratings, subject was instructed to rate the items on a nine-point scale according to their relative ease or difficulty of pronouncing. Low numbers indicate easy to pronounce items; high numbers indicate hard to pronounce items. A free-learning method was used. The lists were presented for alternate study and recall trials, with *S* writing the responses. Each list consisted of nine units to be learned, and each unit consisted of two trigrams, separated by a hyphen. An example of a unit from List 1, the easy list, is WAM-CED; from List 2, the moderately difficult list, KNO-NIQ; and from List 3, the most difficult list, CFL-ZOJ. The 18 trigrams in List 1 had a mean PR of 3.00 with a range from 2.81 to 3.37. For List 2, the mean PR was 5.00 with a range from 4.34 to 5.58. For List 3, the mean PR was 7.00 with a range from 6.37 to 7.60. A total of four study and four recall trials was given. On the study trials, the rate of presentation was 2 seconds with 60 seconds allowed to write down the responses on each recall trial. There was a rest interval of 2 minutes between learning the lists.



## 2. Subjects

Three categories of deaf Ss—A, B, and C—served in the experiment. Category A consisted of 18 Ss who were born deaf with an average hearing loss of 80 db or greater. Category B consisted of 18 Ss who were adventitiously deaf with an average hearing loss of 80 db or greater, and Category C of 18 Ss whose average hearing loss was less than 80 db. All 54 Ss were students selected from the National Technical Institute for the Deaf at Rochester Institute of Technology.

## 3. Procedure

Three groups of 18 Ss each from three categories of deafness learned all three lists, the order of learning the lists being completely counter-balanced among subgroups of six Ss. Three different random orders of each list were made, each order being learned by a different subgroup of six Ss. Each list was presented through a Stoelting memory drum for four free-recall trials, and every S was presented his three lists during a single 30-minute session.

All instructions to the Ss were presented visually in typed cards. No interpreters were employed. To make sure that Ss understood the instructions, they were given a trial list before presented with the experimental lists.

## C. RESULTS AND DISCUSSION

Performance on free-recall trials of paired trigrams for the different categories of deaf subjects is shown in Figure 1 as a function of PR. There, it can be seen that free-recall performance is directly related to rated ease of pronunciation for all categories of deaf Ss. This relation is statistically significant for all three groups. Nonparametric Friedman two-way analyses of variance for groups A, B, and C were  $\chi^2 = 13.19, 13.58, \text{ and } 10.78$ , respectively. With  $df = 2$ , the above  $\chi^2$  values were all significant at the .05 level. Scoring by number of single trigrams produced did not change the conclusion. The results of correct single trigrams of the three lists produced by the groups are shown in Figure 2. Friedman two-way analyses of variance gave the following  $\chi^2$  values: 15.86, 21.36, and 21.11 for groups A, B, and C, respectively. Again, with  $df = 2$ , the  $\chi^2$  values were statistically significant ( $p < .01$ ).

The above findings are of interest, particularly with respect to the category A, deaf Ss. These Ss are profoundly deaf since birth with a hearing loss of at least 80 db or greater. Their audiograms would show no useful

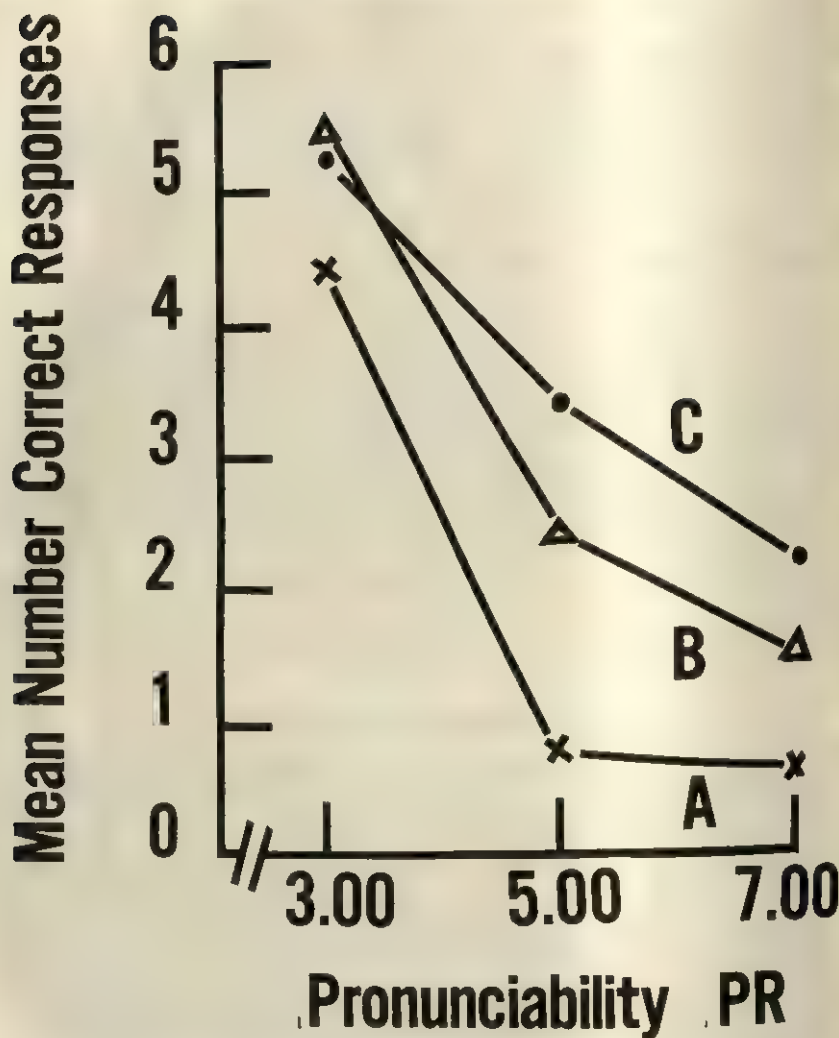


FIGURE 1

FREE RECALL OF TRIGRAM LISTS (EACH UNIT PAIRED) AS A FUNCTION OF RATED PRONUNCIABILITY (PR) AND THE CATEGORIES OF DEAFNESS

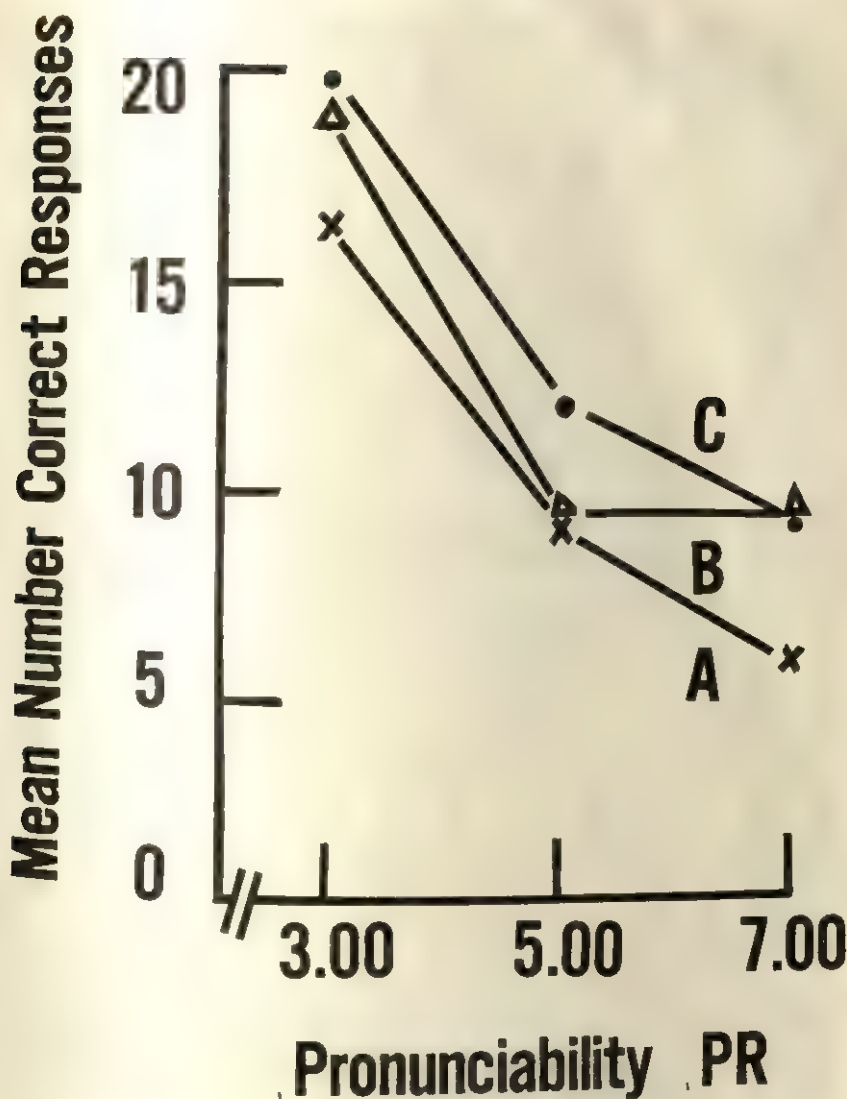


FIGURE 2

FREE RECALL OF TRIGRAM LISTS (SINGLE TRIGRAM AS A UNIT) AS A FUNCTION OF RATED PRONUNCIABILITY (PR) AND THE CATEGORIES OF DEAFNESS

hearing at all across the speech frequency range, and with the best hearing aids available they would "hear" very little speech and perhaps no recognizable speech sounds at all. Their available sensory modes for the intake of verbal information should be crucially different from those of hearing people. The results of their learning of lists with different PRs showed that they, like the hearing subjects (4, 5), telescoped trigram letters into a simplified sound unit for learning. Trigrams that are very difficult to pronounce are not learned easily. In the present experiment, even some profoundly deaf Ss were observed to be pronouncing, finger-spelling, whilst they studied the trigrams presented. Since the Ss also read speech from lips, it is understandable that pronunciability can be a predictor of their verbal learning. They may not experience received auditory stimuli, but they do emit sound, though inaccurately, and are usually taught to do so. The results have implication for teaching of lip reading. That the deaf, with little overt speech, do learn is self-evident.

The free-recall data were also helpful in assessing the comparability of the three groups. The directional differences of rate of learning for the groups AB, BC, and AC were tested by Mann-Whitney  $U$  test for all three lists. On the paired trigram scores for Lists 1 and 3, the  $U$ s for groups AB, BC, and AC were not significant. For List 2, however, group B learned faster than group A ( $U = 109$ ,  $p < .05$ ), and group C also learned faster than A ( $U = 94.5$ ,  $p < .025$ ). Further analysis using the single trigram produced in learning the three lists showed that there were no significant differences in performance among the three groups for Lists 1 and 2. For List 3, group B learned more rapidly than group A ( $U = 99.5$ ,  $p < .025$ ). This is the only  $U$  that was statistically significant for List 3.

The above results seem to indicate that for a list with moderate level of difficulty, the Ss with better hearing learned more rapidly than the profoundly deaf Ss who were born deaf without the advantage of hearing language and may rarely use overt speech. For the very easy list, there was no significant difference in performance among the groups. Probably the profoundly deaf Ss could compensate their lack of hearing by lip reading experience. Words that can be pronounced easily are easily shaped by lips. They articulated the easier trigrams readily, though the pronunciation was not accurate. Again, the better hearing Ss did not do better than the very deaf Ss in the most difficult list. Probably, most of the trigrams in the list could not be easily reduced to simple sound unit and had to be learned letter by letter; hence the variable auditory frequency of experience of sound unit was negligible. When the paired trigram unit was used for scoring, the

moderately difficult level was List 2, but when the single trigram unit was considered, the task requirement for correct responses was easier. The inferior performance of group A as compared to better hearing Ss, group B, was shifted to the learning of List 3. However, group C, though it had a larger mean performance score of 9.22 as compared to 5.83 of group A, did not perform significantly better in the experiment which used a sample size of 18 for each group. The results, despite the small sample size, somewhat support the auditory frequency hypothesis. Auditory frequency of experience is probably a predictor in verbal learning of certain levels of difficult material.

## REFERENCES

1. CONRAD, R. Short-term memory processes in the deaf. *Brit. J. Psychol.*, 1970, **61**, 179-194.
2. CONRAD, R., & RUSH, M. L. On the nature of short-term memory encoding by the deaf. *J. Speech & Hear. Disor.*, 1965, **30**, 336-343.
3. LINDLEY, R. H. Association value, familiarity, and pronunciability ratings as predictors of serial learning. *J. Exper. Psychol.*, 1963, **4**, 347-351.
4. UNDERWOOD, B. J., & POSTMAN, L. Extra experimental sources of interference in forgetting. *Psychol. Rev.*, 1960, **67**, 73-95.
5. UNDERWOOD, B. J., & SCHULZ, R. W. *Meaningfulness and Verbal Learning*. Philadelphia: Lippincott, 1960.

*College of General Studies*  
*Rochester Institute of Technology*  
*One Lomb Memorial Drive*  
*Rochester, New York 14623*





## SELF-IMAGE DISPARITY AND ATTACHMENT TO ETHNIC SUBCULTURE\*<sup>1</sup>

*Yale University*

MARK MCCORMICK AND DAVID BALLA<sup>2</sup>

### SUMMARY

In an effort to investigate the effects of attachment to the dominant culture and of developmental level on the self-concept, 62 male Lebanese-Americans of first, second, or third generation of residence in America completed two forms (questionnaire and adjective checklist) of an instrument designed to measure self-image disparity. A series of questions was used to measure extent of attachment to the Lebanese subculture, and a social competence index was employed as the measure of developmental level. No effects associated with self-image disparity emerged for generation of residence or developmental level. However, subjects who were strongly attached to the Lebanese culture showed less disparity on the questionnaire than those who were detached from it. Some support was provided for the developmental position, since Low Competence Ss made larger numbers of extreme responses on the questionnaire than the High Competence S.

### A. INTRODUCTION

Recently, considerable controversy has developed concerning the relationship between the retention of ethnic cultural heritage and the self-concept (2, 4, 8). Traditionally, in American culture, minority groups have been assimilated into the dominant culture with a consequent abandonment of minority mores and behavior patterns. The value of this initiation has been

\* Received in the Editorial Office on March 5, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This study was supported by Research Grant HD-03008 from the United States Public Health Service, and by the Gunnar Dybwad Award of the National Association for Retarded Children. The authors are grateful to Father Semaan Farris for his cooperation and assistance in the course of this study, and to the members of his parish, St. Anthony's Church, Danbury, Connecticut, for their willingness to participate. Thanks are extended to Elizabeth McCarthy and Christina Orr for their insightful comments and criticisms of an earlier draft of this manuscript. Thanks are extended to Thomas Achenbach, Susan Harter, and Edward Zigler for their critical reading of an earlier draft of this manuscript.

<sup>2</sup> Requests for reprints should be sent to David Balla at the address shown at the end of this article.

so much a part of American tradition that it seemed to remain essentially unquestioned for many years. More recently it has been argued that such a de-emphasis of ethnic tradition is ultimately destructive to the self-concept. It has been suggested that this traditional view of the process of integration into the dominant culture implies that there is a right way to act and a wrong way to act, and the wrong way is the way of the ethnic tradition. Thus, the process of becoming a part of the dominant culture is seen to involve an inevitable deconditioning of ethnic behavior patterns.

It is the purpose of the present study to examine the relationship between attachment to an ethnic tradition and self-image disparity (i.e., the discrepancy between the way an individual views himself and the way he would ideally like to be) in a distinctive cultural group, Lebanese-Americans. The two positions concerning the value of ethnicity would generate antithetical predictions concerning the relationship between attachment to minority culture and self-image disparity. If it is assumed that to become "American" is a positive goal, it would be predicted that people who see themselves as not attached to their ethnic tradition would have less self-image disparity than people who see themselves as highly attached to that tradition. However, if attachment to ethnic tradition is viewed as being related to the positive self-image of minority group members, it would be expected that persons who consider themselves to be highly attached to their ethnic tradition would show less self-image disparity than those who consider themselves to be integrated into the dominant culture.

Although there is no doubt that experiential/psychodynamic factors play a considerable role in self-image disparity, considerable evidence has been presented (1, 5, 11) that any comprehensive understanding of the self-concept must take into account developmental determinants. Thus, a second purpose of the present study was to investigate the developmental determinants of self-image disparity in this sample. It was expected that, independent of attachment to cultural tradition, persons at higher levels of development as measured by a Social Competence Index would show greater disparity than persons at lower developmental levels (1). The rationale for this developmental hypothesis, based on two assumptions, has been presented in detail by Achenbach and Zigler. Firstly, disparity is assumed to increase with increasing capacity to make finer cognitive differentiations. Secondly, persons at higher developmental levels should manifest a greater potential for social guilt; that is to say, they must live up to a greater number of internalized demands.

As in previous studies, four measures were collected for each subject in

order to compute Disparity scores: real self and ideal self, each in two formats—(a) questionnaires in which there were six possible choices and (b) adjective checklists in which there were two possible choices. It was expected that the real-ideal self Disparity would be highest on the questionnaire, where both guilt and cognitive differentiation could operate. Two measures of ethnic attachment were employed in the present study: (a) generation of residence in America, and (b) score on a questionnaire designed to measure ethnic attachment (9). As in the Achenbach and Zigler study, a Social Competence Index was used as a measure of developmental level.

## B. METHOD

### 1. Subjects

A total of 62 male Lebanese-Americans formed the subject pool from which the comparisons to be reported below were taken. The sample was drawn from a single Maronite Roman Catholic parish in Danbury, Connecticut. Of the 75 persons approached, all of whom were identified as having the requisite knowledge of English to complete the questionnaire, 62 participated in the study and completed the instruments.

### 2. Subject Variables

*a. Generation of residence and cultural attachment variables.* All subjects were asked to indicate their generation of residence in the United States. In addition, because simple generation of residence may be an insensitive measure of acculturation, a brief series of questions, adapted from those used by Srole *et al.* (9) and designed to assess the respondent's attachment to Lebanese culture, was administered. The questions were the following: (a) Comparing Lebanese style cooking and regular American style cooking, do you like Lebanese style cooking (i) better, (ii) just as much, or (iii) not as much? (b) During the last year or so did you do anything at all special about *any* of the Lebanese holidays? (c) Are you interested at all in what's going on today in Lebanon? (d) Do you attend Lebanese-American meetings or social affairs often, occasionally, or not at all? (e) Some Lebanese parents feel that it is all right if a son or daughter wants to marry someone who is not Lebanese. Other Lebanese parents don't feel that way at all. How do you feel about it: is it all right or *not* all right?

*b. Social competence scale.* Education, occupation, employment history, and marital status were used as indices of social competence. The general schema was similar to that employed by Phillips and Zigler (7) and by

Zigler and Phillips (10) and has been fully described by these authors. It was impossible to include intelligence of subject as a variable in the current study because of the anonymity guaranteed to the participants. The range of scores for each index was from 0 to 2. The overall Social Competence score for each subject was the mean of the scores obtained on the individual indices. Thus, the final Social Competence score for any *S* could range from 0 to 2. The mean Social Competence scores for the total number of *Ss* were 1.81 and 1.26 for the High and Low Competence groups, respectively, with no overlap between the High and Low Social Competence *Ss*.

### 3. Procedure

Subjects were reached by an initial letter and follow-up phone call. If the subject agreed to participate, the senior author met with the subject, explained the instruments, and arranged for a time to pick them up, within two or three hours if possible. The instruments were arranged in the following order: (a) Social Competence Index and generation of residence questions, (b) cultural attachment questions, (c) real-image questionnaire, (d) real-image checklist, (e) ideal-image questionnaire, (f) ideal-image checklist.

## C. RESULTS

The major dependent measures included (a) Questionnaire Disparity, the number of times the *S* made a response on the ideal-self questionnaire different from the response he made to the same item on the real-self questionnaire; (b) Questionnaire Magnitude Disparity, the degree of difference (answers scored from 0 to 5) between responses to the same statement on the real- and ideal-self questionnaire; and (c) the Adjective Checklist Disparity, the number of times the *S* made a response on the ideal-self list different from the response he made to the same item on the real-self list.

### 1. Disparity by Generation Analysis

From the total subject pool it was possible to match 10 members of each of first, second, and third generation *Ss* for Social Competence. The Disparity scores for these subjects were subjected to a simple effects analysis of variance. No significant effects were revealed by these analyses for any of the measures of Disparity (all  $F_s < 1$ ).

### 2. Generation by Competence and Attachment by Competence Analyses

Since there were not enough members of first generation *Ss* to permit Generation by Competence analyses including first generation *Ss*, 2 (Genera-



tion)  $\times$  2 (Social Competence) analyses of variance were performed for second and third generation Ss. No significant main effects or interaction effects were revealed for any of the measures of Disparity.

In an attempt to provide a more fine-grained measure of attachment to the Lebanese or dominant culture, two groups of persons either Attached to the Lebanese culture or Detached from it were formed with use of the attachment questions mentioned above. The Detached group was comprised of all Ss scoring from 0 through 5 on the question series, and the Attached group had scores from 6 through 10. It was possible to form four groups of 10 persons each of High and Low Attached and High and Low Competence Ss. The relevant Disparity scores for these comparisons are presented in Table 1. A 2 (Attachment)  $\times$  2 (Social Competence) analysis of variance performed on these data revealed significant main effects for Attachment for Questionnaire Disparity ( $F = 10.99$ ,  $df = 1/36$ ,  $p < .005$ ;  $\bar{X}$  Low Attached = 17.2,  $\bar{X}$  High Attached = 11.4) and for Magnitude Disparity ( $F = 7.22$ ,  $df = 1/36$ ,  $p < .05$ ;  $\bar{X}$  Low Attached = 33.0,  $\bar{X}$  High Attached = 20.6) such that the Low Attached Ss had greater Disparity than the High Attached Ss. The main effect for Competence Level and the interaction of Competence Level and Attachment did not reach significance for these measures. On the Checklist Disparity analysis no main effects or interactions reached significance.

### 3. *Further Tests of the Differentiation Hypothesis*

As in previous studies (1, 5) analyses were performed on the number of extreme responses made by the second and third generation High and Low Competence groups on the real-image questionnaire. The results of these analyses are also presented in Table 1. It would be expected that individuals at lower developmental levels would make less differentiated judgments; that is, would tend to choose extreme statements on the questionnaire. A 2 (Attachment)  $\times$  2 (Competence Level) analysis of variance for the extreme response scores revealed a significant main effect for Competence Level such that the Low Competence Ss made significantly larger numbers of extreme responses than the High Competence Ss ( $F = 4.50$ ,  $df = 1/36$ ,  $p < .05$ ;  $\bar{X}$  Low Competence = 15.6,  $\bar{X}$  High Competence = 10.8).

### 4. *Other Analyses*

An effort was made to discover whether the Ss in all three generations had the same perception of the value of the items on the ideal questionnaire and checklist. The number of Ss giving negative responses for each item on the

TABLE 1  
MEAN DISPARITY SCORES AND MEAN NUMBER OF EXTREME RESPONSES OF HIGH AND LOW COMPETENCE, HIGH AND LOW ATTACHED GROUPS

Group	N	Ques- tionnaire Disparity	Ques- tionnaire Magnitude Disparity	Checklist Disparity	Mean number of extreme responses
High Competence, High Attached	10	12.2	21.4	4.5	13.6
Low Competence, High Attached	10	10.5	20.1	4.0	16.3
High Competence, Low Attached	10	17.7	33.0	5.5	8.0
Low Competence, Low Attached	10	16.5	32.8	4.6	14.8

two lists was tabulated. These data were cast into  $3 \times 2$  contingency tables, and Chi squares were computed separately for each item. Of the 60 resulting Chi squares, four reached the 10% level of significance. Since this figure is approximately that expected by chance, it would appear that the items were valued in a similar way among the three generations.

#### D. DISCUSSION

The major finding of the present study would appear to be that, independent of developmental level, subjects who were rated as identified with their cultural tradition showed significantly less self-image disparity than those who were not rated as attached. Empirical support has thus been provided for the position that the process of entry into the dominant culture is purchased at some psychological cost to the self-concept of the members of the minority culture. To the extent that the findings of the present study can be generalized to other cultural minorities, it would appear that expressions of attachment to a cultural tradition have a positive value for personality integration and that educational and other interventional efforts involving minority group members should exhibit an appreciation of their minority heritage.

It is of considerable interest that no association of generation of residence in America and self-image disparity was revealed. This lack of findings associated with the demographic variable of generation of residence, in conjunction with the significant findings associated with a more psychological measure of attachment, underlies the danger of employing such demographic variables as direct indices of psychological experience. Generation of residence in America was only loosely coupled to whatever ex-

periences determined a sense of cultural identity. (Indeed, 67% of the first generation subjects, 55% of the second generation subjects, and 43% of the third generation subjects were found to be highly attached as measured by the attachment questionnaire.)

The results of the present study provided partial confirmation of a developmental component of self-image disparity (1, 5, 6, 11): the Low Competence subjects made more extreme responses on the questionnaire than did the High Competence subjects. This finding is consistent with the hypothesis of increasing cognitive differentiation with increasing development. The finding that Disparity did not vary with Social Competence would appear to indicate that the High Competence subjects did not exhibit the greater guilt that Achenbach and Zigler (1) attributed to an increased capacity to incorporate a greater number of social demands and values. The lack of findings related to guilt suggests that although a person may have a greater capacity for experiencing guilt, guilt itself is not necessarily present. It is intriguing to speculate that the Lebanese cultural tradition does not promote the development of guilt to the same extent that the dominant culture does.

Whatever the speculations concerning the psychological development of guilt in different cultures, it is clear that more fine-grained studies of the interaction of cognitive and experiential factors in determining personality makeup and their relative contributions to such constructs as self-image disparity are in order. Such a need goes beyond idle psychological theorizing. In a recent study by Carpenter and Busse (3) a decrease in real image from the first to the fifth grade in black and white ghetto children was reported. The authors implied that schools failed to build a sense of self-esteem in the children. However, Carpenter and Busse's finding are exactly as would be predicted from the developmental hypothesis. To the developmentalist, the most reasonable inference from the findings would be that formal cognitive development in these ghetto children was proceeding relatively satisfactorily. Thus, in this instance, at least, the separation of cognitive and emotional determinants of self-image would have permitted one to make more relevant comment about the psychological development in our society.

#### REFERENCES

1. ACHENBACH, T., & ZIGLER, E. Social competence and self-image disparity in psychiatric and nonpsychiatric patients. *J. Abn. & Soc. Psychol.*, 1963, **67**, 197-205.
2. AYERS, B. Kelyn and Carolyn. In *Equal Educational Opportunity*. Cambridge, Mass.: Harvard Univ. Press, 1970. Pp. 213-219.

3. CARPENTER, T. R., & BUSSE, T. V. Development of self-concept in Negro and white welfare children. *Child Devel.*, 1969, **40**, 935-939.
4. DAY, N. A. The case for all-black schools. In *Equal Educational Opportunity*. Cambridge, Mass.: Harvard Univ. Press, 1970. Pp. 205-212.
5. KATZ, P., & ZIGLER, E. Self-image disparity: A developmental approach. *J. Personal. & Soc. Psychol.*, 1967, **5**, 186-195.
6. LIGHT, C. S., ZAX, M., & GARDENER, D. H. Relationship of age, sex, and intelligence level to extreme response style. *J. Personal. & Soc. Psychol.*, 1965, **2**, 907-909.
7. PHILLIPS, L., & ZIGLER, E. Social competence: The action-thought parameter and vicariousness in normal and pathological behavior. *J. Abn. & Soc. Psychol.*, 1961, **63**, 137-146.
8. SOMMERS, V. The impact of dual-cultural membership on identity. *Psychiatry*, 1964, **27** (4), 332-344.
9. SROLE, L., LANGNER, T. S., MICHAEL, S. T., OPLER, M. K., & RENNIE, T. A. Mental Health in the Metropolis: The Midtown Manhattan Study. New York: McGraw-Hill, 1962.
10. ZIGLER, E., & PHILLIPS, L. Social competence and the process-reactive distinction in psychopathology. *J. Abn. & Soc. Psychol.*, 1962, **65**, 215-222.
11. ZIGLER, E., BALLA, D., & WATSON, N. Developmental and experiential determinants of self-image disparity in institutionalized and noninstitutionalized retarded and normal children. *J. Personal. & Soc. Psychol.*, 1972, **23**, 81-87.

*Department of Psychology*

*Yale University*

*333 Cedar Street*

*New Haven, Connecticut 06510*

## SELF-PERCEPTION DIFFERENCES AMONG KIBBUTZ AND CITY ADULTS IN ISRAEL AND JEWISH AND NON-JEWISH ADULTS IN THE UNITED STATES\*

*California State University, Northridge*

---

LEO A. PIROJNIKOFF AND ILANA HADAR

---

### SUMMARY

With use of a sample of five kibbutzim and subjects from three cities in Israel, the hypothesis investigated was that the adults in the kibbutz would obtain higher scores than the other subjects in their perceptions of the past and present on the Self-Anchoring Striving Scale. These responses were compared to those of two groups from the United States, Jewish and non-Jewish. The total sample consisted of 160 Ss with 40 Ss in each of the four groups. It was also hypothesized that there would be less discrepancy in the perceptions of the future as compared to the past and present between Israeli and United States Ss. The hypotheses presented were supported by the data.

The kibbutz sample in expressing hopes and aspirations had a significantly higher number of Ss that desired world peace than the other three samples. The Israeli city sample in responding to aspirations had a significantly greater number of responses than the kibbutz for the following categories: having a good job, owning a house, and possessing wealth.

The Israeli city sample compared to the kibbutz had a significantly higher number of responses to a fear of a deterioration in the standard of living. The rank order correlation of the verbal responses indicated that the responses of the American Jewish group were more similar to both Israeli groups than the United States non-Jewish group.

---

### A. INTRODUCTION

In this paper the results of a second and concluding study (6) of cross-cultural differences are reported. The purposes of this investigation were to study the concerns that individuals of different social groups and nations express for themselves and to determine how these groups perceive themselves in satisfactions relating to the past, present, and future.

---

\* Received in the Editorial Office on March 5, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.



The method utilized for this analysis was devised by Cantril (1) and termed the Self-Anchoring Striving Scale. The scale consists of a combination of nonverbal scaling and verbal content analysis.

The samples employed for the study consisted of 40 Ss each from (a) five kibbutzim in Israel, (b) three cities in Israel, (c) American born Jewish Ss from Los Angeles, and (d) American born non-Jewish Ss from the Los Angeles area. The major concern was the difference between the kibbutz and Israeli city samples, with the two United States samples providing information to assess the differences on a broader cultural level.

There are specific characteristics that appear to differentiate the kibbutz from Israeli urban society. One condition suggested by Golan (5) is that in the kibbutz there appears to be less discrepancy between ideology and action. In Israeli urban society, a disparity is generated by differential environmental conditions. These conditions may result in situational pressures which function to prevent the direct expression of ideological convictions and inhibit actions related to these beliefs. There may be sufficient contradictions and ambiguities in the ideology of urban life that limit the possibility of action and distort self-perception. Goffman (4) has stated that a person's self-perception is a function of the demands of various settings.

Another factor leading to divergent perceptions may be the emphasis that communal living has placed on the sharing interactions that occur between people since early childhood. The Kibbutznik can seek satisfaction from all the members in a collective instead of being restricted to his immediate family.

If there are more sources of reinforcement within the kibbutz structure as compared to Israeli urban society, then the kibbutznik is not as vulnerable to manipulation by different environmental situations and has enjoyed a greater sense of security in the past and present. The main hypothesis of this study was that the kibbutz Ss would obtain higher scores of their perceptions of the past and present on the Self-Anchoring Scale as compared to non-kibbutz Ss in Israel and the two groups from the United States.

The future has aspirations and goals that are interwoven with the preservation of the kibbutz and Israel in the present political situation in the Middle East. Since the future is less distinct than the past or present, it was also hypothesized that there would be less discrepancy in the perception of the future as compared to the past and present between the Israeli and United States Ss. Another purpose of the investigation was to explore how the four groups differed in verbal aspirations and fears for self.

## B. METHOD

The total sample consisted of 160 Ss. The characteristics of the sample were presented in a previous study (6). The Ss in the four groups were administered the Self-Anchoring Striving Scale. The Ss were first presented with a questionnaire that contained two items dealing with their hopes and fears for self. After completing their answers to these two questions, the Ss were asked to turn the page and complete the nonverbal scale. This scale was in the form of a ladder with steps numbering from zero to 10. The Ss were asked where they stood on the ladder today, with the top being the best life, the bottom the worst life; where they stood in the past; and where they thought they would stand in the future. At the end of each question there was a blank in which the Ss were to write their numerical estimates of each of the questions.

## C. RESULTS

Table 1 contains the means and standard deviations for the past, present, and future categories on the Self-Anchoring Striving Ladder for each of the four samples.

An analysis of variance was performed on each of the three classifications. The  $F$ s obtained for the past and present categories were significant ( $F = 5.86$ ,  $df = 3/156$ ,  $p < .01$ ;  $F = 4.07$ ,  $df = 3/156$ ,  $p < .01$ ). The  $F$  for the future was not significant. In all categories, the kibbutz scored higher than the other samples. Scheffé's (2) test indicated that the kibbutz sample was significantly different from the other three groups and that the significant  $F$ s were all determined by this difference.

TABLE 1  
MEANS AND STANDARD DEVIATIONS FOR THE FOUR GROUPS ON THE THREE CATEGORIES  
OF THE SELF-ANCHORING STRIVING LADDER

Group	Past	Present	Future
Kibbutz	6.20 $\pm$ 2.11	7.25 $\pm$ 1.50	8.52 $\pm$ 1.52
Israeli-city	5.00 $\pm$ 2.15	6.22 $\pm$ 1.56	8.12 $\pm$ 1.09
U. S. Jewish	4.75 $\pm$ 2.28	6.08 $\pm$ 2.5	8.05 $\pm$ 1.41
U. S. non-Jewish	4.25 $\pm$ 2.10	6.22 $\pm$ 1.49	8.22 $\pm$ 1.87

Personal aspirations were coded according to the 14 categories devised by Cantril and are presented in Table 2. The kibbutz had a significantly greater number of individuals desire world peace than the other three samples. The Israeli city sample had a significantly larger frequency of indi-

TABLE 2  
GROUP PERCENTAGES FOR PERSONAL HOPES AND ASPIRATIONS ON  
THE SELF-ANCHORING STRIVING SCALE

Categories	Israel		United States	
	Kibbutz	City	Jewish	Non-Jewish
Happy family life	75.0	72.5	62.0	60.0
World peace	40.0	22.5	15.0	17.5
Good job	22.5	37.5	30.0	42.5
Self-development or improvement	20.0	22.5	27.5	25.0
Improved or decent standard of living	20.0	17.5	30.0	15.0
Opportunities for children	12.5	20.0	7.5	5.0
Achieve sense of own personal worth	12.5	5.0	10.0	10.0
Acceptance by others	10.0	10.0	15.0	25.0
Social justice	7.5	.0	12.5	10.0
Concern for relatives	5.0	10.0	7.5	15.0
Emotional stability	5.0	.0	5.0	10.0
Own house	2.5	22.5	10.0	5.0
Have wealth	2.5	22.5	7.5	5.0
Desire to be useful to others	.0	.0	10.0	17.5

viduals than the kibbutz responding on aspirations for the following categories: having a good job, owning a house, and possessing wealth.

Personal worries and fears were coded according to the 12 categories provided by Cantril and are presented in Table 3. The kibbutz had a significantly larger number of responses to fears and worries concerned with both war and an unhappy family life than the Israeli city sample. The

TABLE 3  
GROUP PERCENTAGES FOR PERSONAL WORRIES AND FEARS ON  
THE SELF-ANCHORING STRIVING SCALE

Categories	Israel		United States	
	Kibbutz	City	Jewish	Non-Jewish
Unhappy family life	55.0	35.0	37.5	7.5
War	47.5	22.5	25.0	7.5
No self-development or improvement	20.0	25.0	32.5	25.0
Emotional instability	20.0	7.5	7.5	10.0
Deterioration in standard of living	12.5	45.0	27.5	17.5
Poor job	12.5	20.0	32.5	27.5
Not accepted by others	12.5	7.5	37.5	32.5
Separation from relatives	10.0	5.0	7.5	17.5
Status quo	7.5	12.5	7.5	12.5
Unemployment	7.5	7.5	.0	2.5
No sense of personal worth	7.5	5.0	5.0	20.0
Not to be useful to others	.0	.0	2.5	12.5

Israeli city sample had a significantly higher frequency than the kibbutz on the possible deterioration in the standard of living.

Six rank order correlations were obtained on each of the tables of aspirations and hopes and worries and fears. The correlations between the three Jewish groups ranged from .58 to .79. When the U.S. non-Jewish group was introduced into these comparisons, the correlations diminished to a range of from  $-.07$  to  $.44$ .

#### D. DISCUSSION

The kibbutz sample obtained significantly higher scores than the other three samples on the past and present categories of the Self-Anchoring Striving Scale. Gerson (3) suggested that the self-concept is a result of social learning and that social reinforcement is one of the most important determinants for the self-concept. In the realm of economics and politics, the kibbutznik tends to be more secure and experiences little worry for the family support or the children's education. Since the environment presents the kibbutz Ss with numerous possibilities of positive reinforcement that may be more attainable, immediate, and frequent, their feelings of satisfaction and security may tend to be more dominant than for the other three groups. Consequently, for the kibbutz Ss, there may be a greater feeling of success and attainment which was reflected on both the past and present categories of the Self-Anchoring Striving Scale. All the hypotheses presented were supported by the results of this study.

The attitudes against war and for peace and a better world are generated from the kibbutz's general thesis that the members are responsible in all areas of living for one another's welfare, as well as the nation as a whole.

The Israeli city sample has been characterized many times by different writers as the affluent American society. This society can be perceived as one in a continual state of flux reflecting economic and social growth. It would therefore appear that the verbal responses for the desire of material goods support this view.

In the rank order correlation of verbal responses, the American Jewish group responses were more similar to both Israeli groups than the United States non-Jewish group. This may indicate a greater similarity of interests in values among the three Jewish groups as opposed to the United States non-Jewish group.

#### REFERENCES

1. CANTRIL, H., & FREE, L. Hopes and fears for self and country: The Self-Anchoring Striving Scale in cross-cultural research. *Amer. Behav. Sci.*, 1962, 6, Monog. Suppl. 2.

2. EDWARDS, A. *Statistical Methods*. New York: Holt, Rinehart & Winston, 1967.
3. GERSON, M. *On Education and the Family in the Kibbutz*. Merchaviah, Israel: Hakibbutz Haartzi, 1968.
4. GOFFMAN, E. *The Presentation of Self in Everyday Life*. New York: Doubleday, 1959.
5. GOLAN, S. *Kibbutz Issues*. Merchaviah, Israel: Hakibbutz Haartzi, 1961.
6. PIROJNIKOFF, L., HADAR, I., & HADAR, A. Dogmatism and social distance: A cross-cultural study. *J. Soc. Psychol.*, 1971, 85, 187-193.

*Department of Psychology*  
*California State University, Northridge*  
*Northridge, California 91324*



LYSERGIC ACID DIETHYLAMIDE (LSD 25): XXXX. EFFECT OF  
pH ON TRANSPORT OF METHYSERGIDE AND LSD 25  
ACROSS GILL MEMBRANE\*<sup>1</sup>

*South Oaks Research Foundation, Amityville, New York*

---

H. H. GETTNER, P. A. CARONE, AND H. A. ABRAMSON

---

SUMMARY

The effect of pH on the transport of lysergic acid diethylamide (LSD 25) and of methysergide has been studied in both Siamese fighting fish and goldfish. Two methods have been employed to study the transport of these drugs. In one method, the fish were immersed in a solution of the drug, thus providing an essentially unlimited supply of the drug in the outside liquid during the course of the experiment. In the second method, the fish were dipped for a much shorter period and then observed after the excess drug had been washed off. It has been found that the pH of the outside liquid, in both methods, is a major factor controlling the transport of LSD 25 and methysergide across the gills. Between pH 5.0 and pH 5.7, the surfacing reaction of the fish is diminished, with maximum blocking occurring at pH 5.0, the blocking effect diminishing to zero at pH 5.8. The data are discussed in connection with the permeability of the blood brain barrier of the nervous system to psychotomimetic drugs.

---

A. INTRODUCTION

In preliminary experiments (2), histidine, glutamic, and dl-aspartic acids (at pH 6.0) seemed to partially block the transfer of LSD 25 across the gill of the Siamese fighting fish as determined by the surfacing reaction. For a randomly selected group of other simple amino acids, the pH had to be lowered to 5.0-5.6 to obtain the blocking effect. The study of the effect of pH of the liquid bathing the gills on gill permeability therefore seemed pertinent. For this latter study we used goldfish rather than Siamese fighting fish. LSD 25 and the butanolamide of d-lysergic acid (methysergide) were employed in these studies.

---

\* Received in the Editorial Office on March 5, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The authors' thanks are due Sandoz Pharmaceuticals for generous donations of methysergide used in these experiments.

The word transport here is used to indicate movement of compounds studied through a membrane without reference to chemical or physical processes occurring within the membrane itself.

### B. METHODS

Two methods were employed to study the effect of pH on the passage of LSD 25 and methysergide through the gills of goldfish. In the first method, 10 goldfish, one to one-and one half inches in length are placed into liter graduate cylinders (height = 35 cm) containing one liter of test solution, thus providing a large and fairly constant quantity of drug throughout the experiment. A slow, steady stream of air is bubbled through each cylinder containing the goldfish. Similar techniques were used previously (2) for Siamese fighting fish, but without the air stream. For each reading of the number of fish surfacing when air bubbling is used, the air is turned off to avoid having the fish brought to the surface by the stream of air bubbles. The number of nose-up, tail-down fish surfacing is observed every 10 minutes.

In the second method, five goldfish, one to one-and-one half inches in length, are dipped into 100 ml of the test solution for five minutes and then washed successively for 10 seconds in four beakers, each containing 100 ml of distilled water. The fish are then placed into liter cylinders containing one liter of distilled water (no drug) with air bubbling through. The number of fish surfacing in the nose-up, tail-down position is noted at 10-minute intervals, as in the first method. From a methodological viewpoint, the dipping method corresponds to a rapid intravenous infusion of the drug, the transfer occurring across the gill membranes.

### C. RESULTS

In Figure 1 are the results of experiments ( $n = 290$ ) reported previously on Siamese fighting fish after immersion in solutions of LSD 25 for different initial concentrations (2). The solid black lines in Figure 1 show the standard surfacing reaction when the fish are constantly immersed in solutions of LSD 25. Compare the surfacing reaction when the fish are briefly dipped into solutions of LSD 25 for 30 seconds. The concentrations of LSD 25 varied from 5 to 50 mcg/ml. The die-away curves are very different from the standard solid line persistent immersion curves. The die-away curves indicate in our opinion that the LSD 25 diffused primarily through the gills when the fish were dipped into LSD 25. The LSD 25 is probably destroyed by metabolic processes, rather than excreted. Dipping is probably equivalent

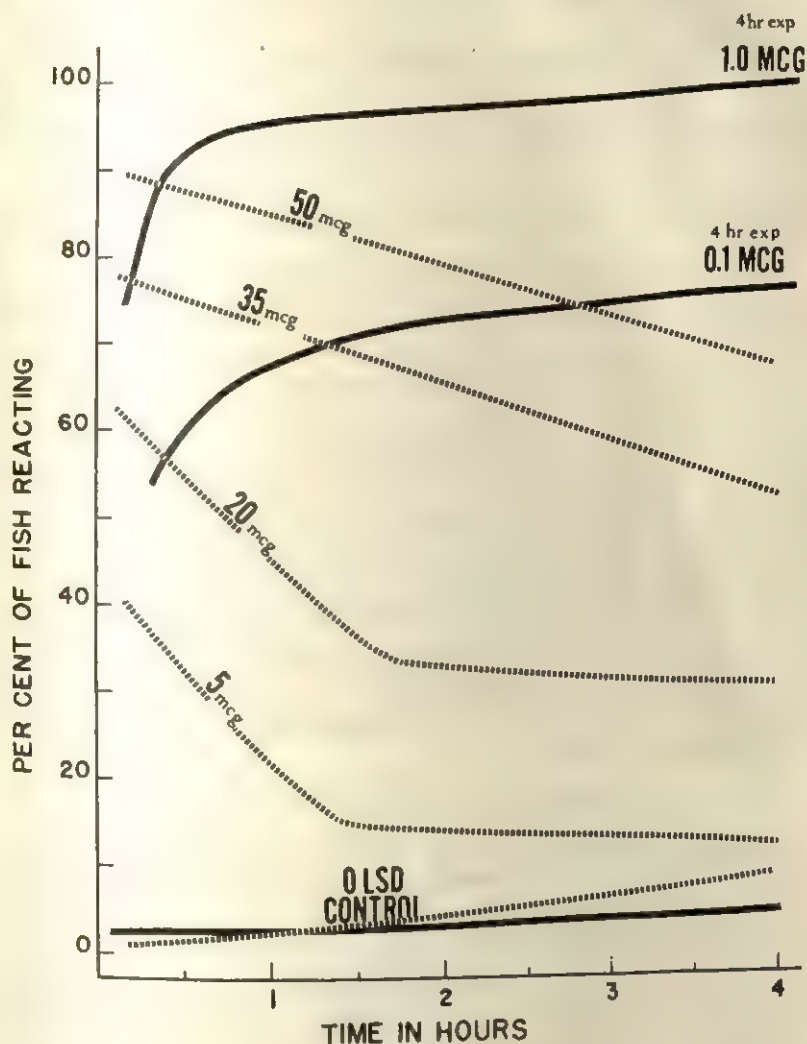


FIGURE 1

EFFECT OF DIPPING FISH INTO LSD SOLUTIONS OF VARYING CONCENTRATIONS  
(TIME—30 SECONDS)

The solid black lines show the surfacing reaction of Siamese fighting fish when immersed in solutions of LSD 25. The dotted lines represent the surfacing reaction when fish are dipped for 30 seconds into solutions of LSD 25 varying in concentration from 5 to 50 mcg/ml.

to an injection. Die-away curves are obtained both after injection and dipping. Similar die-away curves were obtained with goldfish by dipping for 30 seconds into solutions of LSD 25 or higher concentrations of methysergide. In connection with our belief that transport is primarily through the gills, the Information Bulletin of Sandoz, Ltd., Basle, Switzerland, points out that when methane sulfonate of ethyl m-aminobenzoate is used as an anesthetic, the solution is active if it is sprayed on the gills of large fish; e.g., sharks or rays. Given orally to mammals, the anesthetic produces no effects.

### 1. *The Effect of Goldfish on the pH of the Medium*

HCl solution was added to distilled water to bring the pH to 5.0. Ten goldfish swimming freely, as indicated under Methods, within 10 minutes changed the pH to 5.4-5.5. At the end of one hour, the pH had risen to 5.8 ( $n = 30$ ). Similarly, in experiments using LSD 25 or methysergide, when the pH was 5.0 at the start of the experiment in an unbuffered solution, it rapidly rose, finally reaching pH 5.7 at the end of one hour. This is probably due to secretion of basic material. When the pH initially was 6.0, there was little if any pH change at the end of one hour. This rise in pH complicated the experimental procedure somewhat, since buffers were required. With .01 M sodium phosphate buffer, no appreciable pH change occurred during an experiment of one to two hours. The effect of LSD 25 and methysergide on the surfacing reaction of the goldfish was then readily studied. The presence of LSD 25 or methysergide in our system does not materially affect the rise in pH.

### 2. *Is pH 5.0 a Suitable Medium for Goldfish?*

According to authorities (1) and our own experience, water at pH 5.0 adjusted with HCl or with phosphate buffer and HCl does not damage the fish. Fish kept in solutions of pH 5.0 to pH 7.0 for hours do not show gross evidence of damage. To demonstrate the rapid and reversible change in transport of LSD 25 as measured by the surfacing reaction as a function of pH, see Figure 2. In Figure 2, the solid black curve at the bottom demonstrates the slow transport (surfacing reaction) of LSD 25 across the gill membrane when base secreted by the fish is neutralized by small additions of HCl to maintain the pH near 5.0. At the end of one hour, the pH was rapidly changed to 5.7 by addition of NaOH. Note in Figure 2 the rapid surfacing of the fish corresponding to the surfacing produced by the remaining reservoir of LSD 25 in the outside liquid. The concentration of

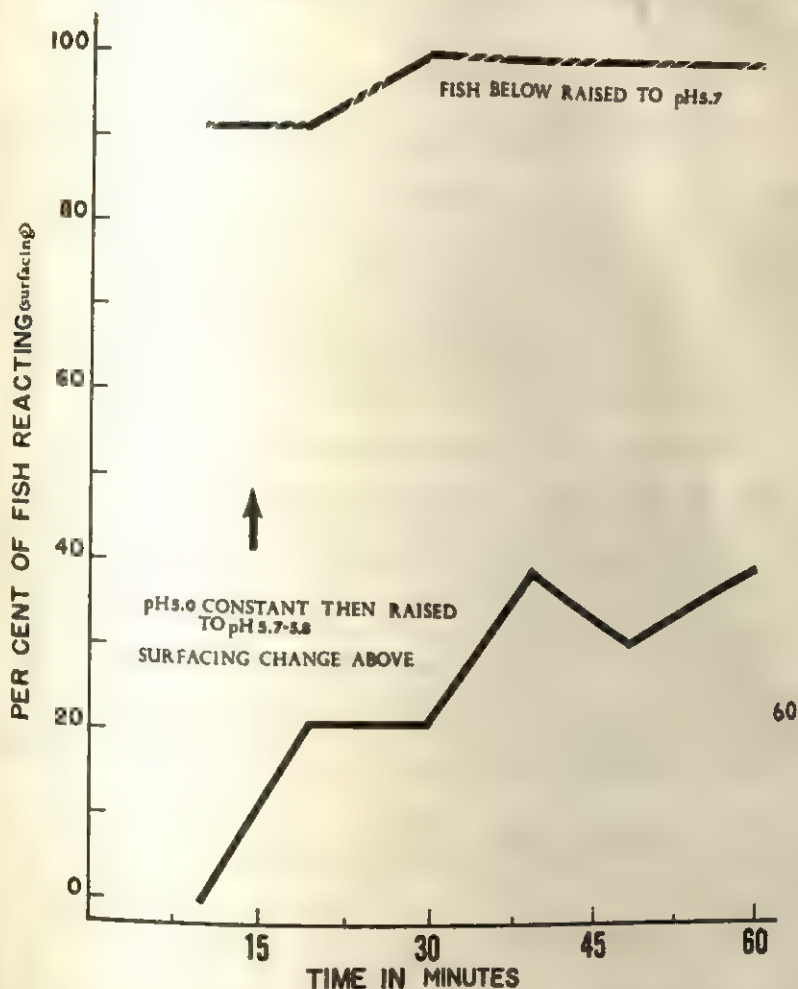


FIGURE 2

## EFFECT OF SUDDEN RISE IN pH ON SURFACING OF FISH

The solid black curve at the bottom demonstrates the slow transport of LSD 25 across the gill membrane when base secreted by the fish is neutralized by small additions of HCl to maintain the pH near 5.0. At the end of one hour the pH was rapidly changed to about 5.7 by addition of NaOH. Note the rapid surfacing of the fish corresponding to the surfacing produced by the remaining reservoir of LSD 25 in the outside liquid.



LSD 25 at pH 5.0 was apparently not materially affected during the experiment.

### 3. The Cutoff pH of Blocking in Simple Systems Using HCl

In order to approximate the pH at which there seemed to be little effect of pH in simple systems, HCl was added initially to the LSD 25 or methysergide solution to adjust to the desired pH, and observations were made for one hour. With HCl between pH 5.5 and 5.8, it appeared that the effect of pH becomes negligible at approximately pH 5.8 for .5 mcg/ml of LSD 25. But below pH 5.7-5.8, the transport of LSD 25 across the gill membrane is definitely decreased. The final surfacing reaction, in the absence of buffers, however, is about the same after one hour because of two factors: increase of pH and a large reservoir of LSD 25.

### 4. Comparison of LSD 25 and Methysergide by the Dipping Method

With use of the dipping method, the surfacing of goldfish dipped for five minutes in .01 M sodium phosphate buffer containing 3.5 mcg LSD 25/ml at pH 5.0 was compared with those dipped into a similar LSD 25 buffer solution at pH 6.0. Since almost 100% of those dipped at pH 6.0 surfaced rapidly, the blocking reaction is calculated from the number of fish that do not surface at pH 5.0. That is, the percent of fish that do not surface is the percent of blocking. The data at pH 6.0 should be considered as 0% of blocking and are therefore not included in Table 1. Similar experiments were run with methysergide at a concentration of 50 mcg/ml. The data in Table

TABLE 1  
BLOCKING OF SURFACING REACTION OF GOLDFISH AT pH 5.0<sup>a</sup>

Goldfish <i>N</i>	Concentration	Dipping time (minutes)	% blocking at time in minutes					60
			10	20	30	40	50	
<i>LSD 25 solution</i>								
40	3.5 mcg/ml	5	35	42	56	64	60	67
10	1.75 mcg/ml	5	44	56	57	71	80	100
10	3.5 mcg/ml	2.5	60	50	66	37	25	37
<i>Methysergide solution</i>								
35	.5 mg/ml	5	26	31	37	40	40	46
10	.25 mg/ml	5	20	50	83	80	100	75
10	.5 mg/ml	2.5	38	29	43	29	20	34

<sup>a</sup> Buffer: .01 M sodium phosphate solution. The fish were dipped into 100 ml of (a) LSD 25 solution in the first set of tests and (b) methysergide solution in the second set of tests; and then they were washed four times successively for 10-second periods in 100 ml of distilled water.

1 show unequivocally that blocking of the surfacing reaction occurs at pH 5.0 by use of the second method, that of dipping for a brief period.

Since the amount of LSD 25 and methysergide that goes through the gills while in contact with LSD 25 or methysergide as determined by the surfacing reaction is less at pH 5.0 than pH 6.0, the effects at pH 5.0 are like those of a lower concentration of the drug at pH 6.0.

#### D. DISCUSSION

Our interest in the transfer of LSD 25 and methysergide across the gill membrane was originally engendered by earlier observations that, under certain conditions, beef brain and other tissue extracts blocked this transfer. Later observations indicated that the pH of the medium bathing the fish, and therefore the gills, markedly affected the percentage of fish surfacing when LSD 25 or methysergide was in the medium. We assume here that the effect of pH reported is primarily due to the effect on the gill transport system, and not due to the ionization or charge of the drugs sharply changing between pH 5.0 and 5.7, or to the overall effect of these pH changes on the skin reflexly influencing the gas content of the swim bladder.<sup>2</sup> The fact that the pH of the medium may influence absorption through living membrane is well borne out by data of Travell (5) on the stomach of the dog and cat, with use of alkaloids. In animals with ligated cardia and pylorus, alkaloids are not absorbed to any extent from the stomach when the reaction of the gastric juice is strongly acid. If the gastric juice is rendered alkaline, however, alkaloids are rapidly absorbed from the ligated stomach. It is of interest that absorption slows at about pH 5.0 and is fairly rapid at pH 6.0.

Experiments on the transport of LSD 25 and its congeners and derivatives are not without clinical interest. The permeability of the blood brain barrier to LSD 25 is low. It is of importance to determine all those conditions that lead to the absorption of drugs from the blood stream into the central nervous system. Without data on living tissues, theoretical extrapolations for the psychotomimetic drugs are rarely valid. Another factor to be considered is that the pH at the surface of a tissue is in all likelihood different from the pH of the medium itself. In the case of the gill, ammonia according to Smith (4), is excreted. When absorption occurs, the living membrane is a dynamic system, and it is not surprising that the value of the pH of the medium influences the dynamics at the cell's surfaces.

<sup>2</sup> We do not have available the dissociation curves of these compounds between pH 5.0 and 5.7. These will be obtained, if possible, and considered, if necessary.

In the case of the gill, ammonia is excreted, according to Smith. However, in our opinion, it is difficult to visualize ammonium hydroxide being responsible for the change. We feel it is more likely that the gill would act very much like the pancreas and that the increase in pH of the outside liquid from pH 5.0 to pH 5.8 is due to the excretion of bicarbonate. This is discussed at length by Lehniger (3).

#### REFERENCES

1. BROWN, M. A., *Ed.* The Physiology of Fishes (Vol. II, Behavior). New York: Academic Press, 1957. P. 411.
2. GETTNER, H. H., ROLO, A., & ABRAMSON, H. A. Lysergic acid diethylamide (LSD 25): XXXVI. Comparison of effect of methysergide (UML 491) on goldfish and Siamese fighting fish. *J. of Psychol.*, 1965, **61**, 87-92.
3. LEHNIGER, A. L. Biochemistry. New York: Worth, 1970.
4. SMITH, H. Fish Physiology (Vol. I). New York: Academic Press, 1969. P. 314.
5. TRAVELL, J. The influence of the hydrogen ion concentration on the absorption of alkaloids from the stomach. *J. Pharmacol. & Exper. Ther.*, 1940, **69**, 21-33.

*South Oaks Foundation, Inc.*

*Sunrise Highway*

*Amityville, New York 11701*

## THE EFFECT OF EXPERIMENTER ABSENCE AND RESPONSE DELAY ON NONREINFORCED IMITATION\*<sup>1</sup>

*Bureau of Child Research, University of Kansas*

PAUL M. SMEETS AND SEBASTIAN STRIEFEL

### SUMMARY

Present research attempted to demonstrate the effect of response delay and the presence or absence of the experimenter on the performance of nonreinforced imitative behaviors in retarded children. A multiple baseline design, counterbalanced for conditions, was used. The results indicated that the degree of control exercised by the experimenter's presence or absence was partially determined by the response delay. Other variables that may have affected the imitative performance were (a) other nonimitative behaviors displayed by the subject(s); (b) additional stimuli generated by the experimenter; and (c) the subject's pre-experimental history with the experimenter.

### A. INTRODUCTION

For several years, research on imitation has focused on variables controlling the development and maintenance of nonreinforced imitation. It has been demonstrated that the performance of nonreinforced imitative behaviors may be a function of the reinforcement of other imitative (1, 2, 4, 12, 13) and nonimitative behaviors (14). The strength of nonreinforced imitative behavior has also been shown to be highly controlled by various setting events. A setting event (10) refers to antecedent and concurrent stimuli in the environment which modify the probability of occurrence of various subsequent responses. Of such setting events as the type of instruction (5, 6, 16, 17, 18, 21), the model's sex (21), and the subject's pre-imitation history with the model (3, 8), the presence of the experimenter (15) has received considerable attention. Peterson and Whitehurst (15) observed that, for all three

\* Received in the Editorial Office on March 12, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This investigation was partially supported by Parsons State Hospital and Training Center, the Training Grant HD-00183 from the National Institute of Child Health and Human Development to the Kansas Center for Research in Mental Retardation and Human Development, and NICHD Grant HD-00870. Reprints may be obtained from Paul M. Smeets at the address shown at the end of this article.

subjects used (Experiment I), the strength of reinforced and nonreinforced imitative behaviors decreased by 50 percent or more under the condition in which the experimenter was absent during the emission of the response. Similar results were obtained (Experiment II) when none of the imitative responses was reinforced. Under the condition in which the experimenter was absent during response emission, the imitative behaviors of the four subjects used dropped by 88, 40, 22, and 15 percent, respectively. Peterson and Whitehurst (15) concluded that the presence or absence of the experimenter exercised moderate to strong control over the performance of nonreinforced imitative behavior. However, it is not clear whether the obtained results were affected by other variables different from but closely associated with the experimenter's presence or absence, such as response delay. Recent evidence (11) indicates that the imitative performance of retarded and normal children will decrease as a function of delay between stimulus presentation and opportunity to perform. The present study was an attempt to evaluate the effect of response delay on the performance of nonreinforced imitative behaviors during conditions in which the experimenter is present or absent.

## B. METHOD

### 1. *Subjects and Apparatus*

Five boys and one girl served as subjects. Subjects were selected on the basis of IQ and availability. The ages ranged from 10.8 to 13.2 years (mean age: 12.1 yrs). The IQs (Peabody Picture Vocabulary Test) varied from 52 to 78 (mean IQ: 65.5).

An adult male served as experimenter. The study was conducted in an experimental room with a one-way mirror which was covered with paper. The paper had an opening large enough for outside observations. Furnishings included two chairs and a table with a stopwatch scoring sheets, and cue cards. The subject and the experimenter were seated facing each other. Near the experimenter's chair was a foot pedal to activate a buzzer. The buzzer was programmed by a timer located in the adjacent observation room.

### 2. *Procedure*

Subjects were seen once a day, five times a week. Because of time limitations, some subjects were seen twice a day during the terminal phase of the study. The first session included a procedure to train the subject to imitate the experimenter only after a buzzer sounded. After the sound of the buzzer had been demonstrated to the subject, the experimenter modeled a series of



verbal items, one at a time. The delay period between the modeling of the behavior and the sound of the buzzer was randomly either .5 or 8 sec. This training was continued until the subject made five correct responses in succession. The instruction, "Do this when the buzzer goes," was repeated at the beginning of each trial during the entire study. Each session consisted of 15 trials. A trial consisted of the instruction, the modeling of a behavior, and the opportunity for the subject to imitate. The total duration of each trial was 25 sec. None of the subject's responses was followed by observable consequences other than scoring of the response. The modeled behaviors were as follows: (a) fold arms, (b) wave hand, (c) clap hands, (d) hand on head, (e) raise both arms, (f) one arm forward, (g) nod yes, (h) raise one arm, (i) stand up, (j) both arms sideways, (k) walk around chair, (l) slap leg, (m) both arms forward, (n) swing torso sideways, (o) cover both ears. The order of stimulus presentations was randomized each session.

The experiment included the following conditions:

a. *Experimenter Present—no delay.* During this condition, the experimenter would stay in the room for the entire session. After modeling the behavior, the experimenter would press the foot pedal which would start the timer, and the buzzer would sound within a .5 sec period.

b. *Experimenter Present—8 sec delay.* This procedure was identical to the previous one with the exception that the delay between modeling of the behavior and the sound of the buzzer was extended to 8 sec.

c. *Experimenter Absent—8 sec delay.* The experimenter would leave the room immediately after having modeled a behavior. The 8 sec interval between modeling of the behavior and the sounding of the buzzer allowed the experimenter to be out of the room for one to one-half seconds before the buzzer sounded. The experimenter stayed out of the room an additional 8 sec after the buzzer.

A multiple baseline, counterbalanced for conditions, was used. The subjects were divided into three pairs. The same experimental conditions were used for all pairs but in a different sequence. Each pair was exposed to a different condition initially. A multiple baseline was collected for subjects of the same pair so that they were exposed to the same sequence of conditions at different points of time.

After each session, the subject was given a candy bar or five cents, in a different part of the building, by a person not involved in the study. Thus, money and candy bars were contingent only on having left the experimental room.

On 105 different occasions during the conditions in which the experimenter

was in the room, the subjects' responses were scored independently by both the experimenter and an observer seated behind the one-way mirror. A response was considered correct if (a) its topography did not deviate from predefined behavior characteristics, and (b) it was initiated after the buzzer sounded but before the beginning of a new trial or before the experimenter re-entered the room.

### C. RESULTS

Reliability on rater scoring ranged from 86 to 100 percent, with a mean of 98 percent. Reliability was computed by dividing the number of trials where both the experimenter and observer agreed that correct imitation had occurred by the total number of trials per session.

Figures 1, 2, and 3 show the effects of the various procedures on the performance of imitative behavior. The figures indicate that except for subject 1, who continued to imitate at a high and stable rate under all three conditions, the rate of correct imitation was a function of the types of procedures used, regardless of their sequential order. Under the condition of "Experimenter Present—no delay," subjects 2, 3, 4, 5, and 6 performed more correct imitations than in either of the other conditions. In addition, the performance of subjects 2, 4, 5, and 6 was more stable under this condition than under either condition, with 8 sec delay. Under the two conditions of response delay, the imitative performance of subjects 2, 3, 5, and 6 was higher when the experimenter was present than when he was absent. Except for the initial sessions of the "Experimenter Present—8 sec delay" condition, subject 4's overall imitative performance during this condition was not much different from that observed during the condition in which the experimenter was absent. A marked decrease in imitative responding over sessions was observed for subjects 3 and 4 during the "Experimenter Present—8 sec delay" condition. This was the initial condition for subjects 3 and 4. Both subjects exhibited a high rate of imitative responding during the initial session(s), which rapidly declined during the following sessions. A similar, but less dramatic decrease was observed for subject 5 under this condition. In contrast, subjects whose performance was markedly affected by the absence of the experimenter (subjects 2, 3, 4, and 5) showed a low performance rate as of the first session of that condition.

Subject 3 showed no great difference in overall imitative performance between the two initial conditions ("Experimenter Present—8 sec delay" and "Experimenter Present—no delay"). However, imitation was continuing to decrease under the condition "Experimenter Present—8 sec delay" when

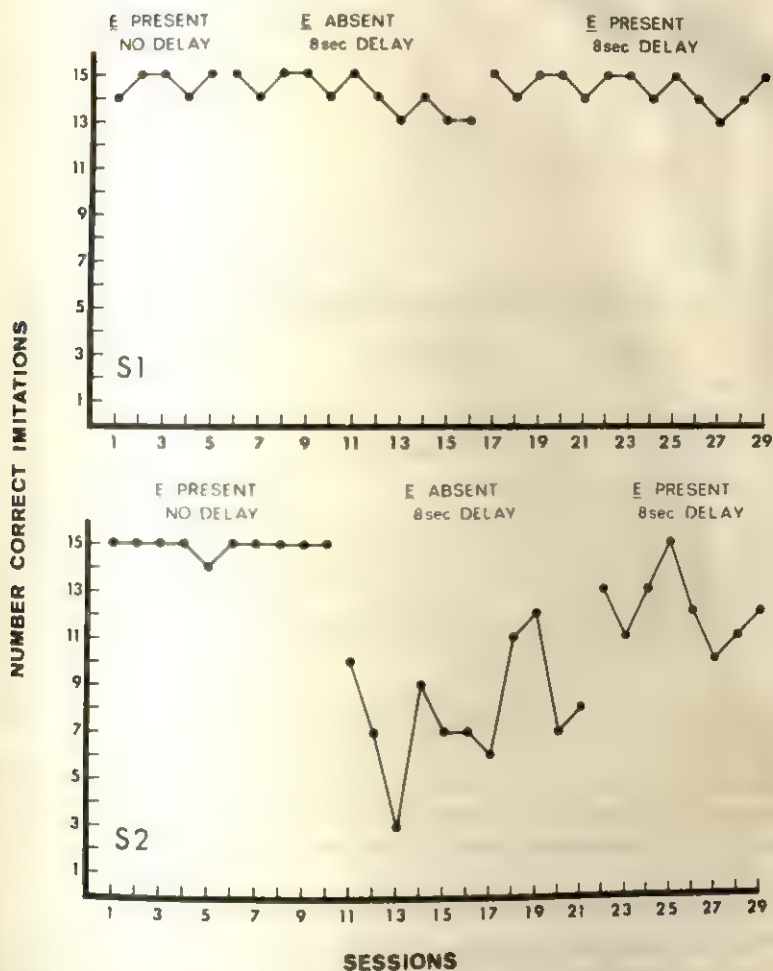


FIGURE 1

NUMBER OF NONREINFORCED IMITATIONS UNDER EACH CONDITION FOR SUBJECTS 1 AND 2

this condition was terminated. A performance increase was observed during the last sessions of the "Experimenter Present—no delay" condition. In order to explore whether the subject's performance under these conditions might have stabilized at different levels if the conditions had been continued for more sessions, subject 3 was re-exposed to these two conditions after he had been exposed to all the experimental conditions. The data indicate that subject 3's imitative performance was not completely replicated

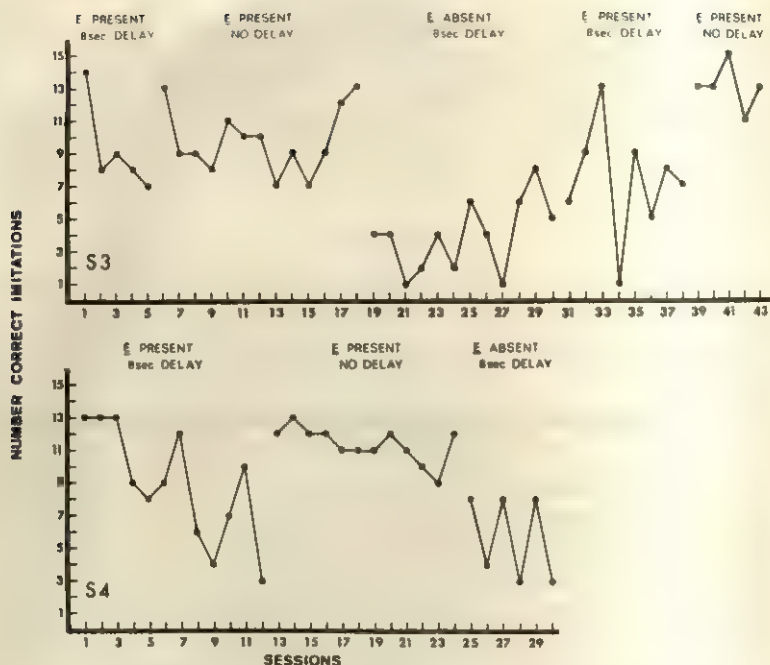


FIGURE 2

NUMBER OF NONREINFORCED IMITATIONS UNDER EACH CONDITION FOR SUBJECTS 3 AND 4

during the reversal. A higher and more stable rate of imitative responding was observed during the "Experimenter Present—no delay" condition than during the "Experimenter Present—8 sec delay" condition.

#### D. DISCUSSION

When the experimenter remained in the room, presenting the subject the instruction, "Do this," giving him the opportunity to respond immediately, modeled behaviors were imitated at a high frequency although they were not reinforced. Such findings are in accordance with those of previous studies (5, 6, 15, 17, 18, 19), indicating the strong social controls exercised by the presence of the model and the instruction "Do this."

The decrease in imitative performance under the 8 sec delay conditions may be related to the inadequacy of short-term memory commonly observed in retarded individuals (7, 9, 20). However, the increased variability in performance under these conditions suggests that during the 8 sec delay, the performance of imitative behaviors may have been subject to other variables.

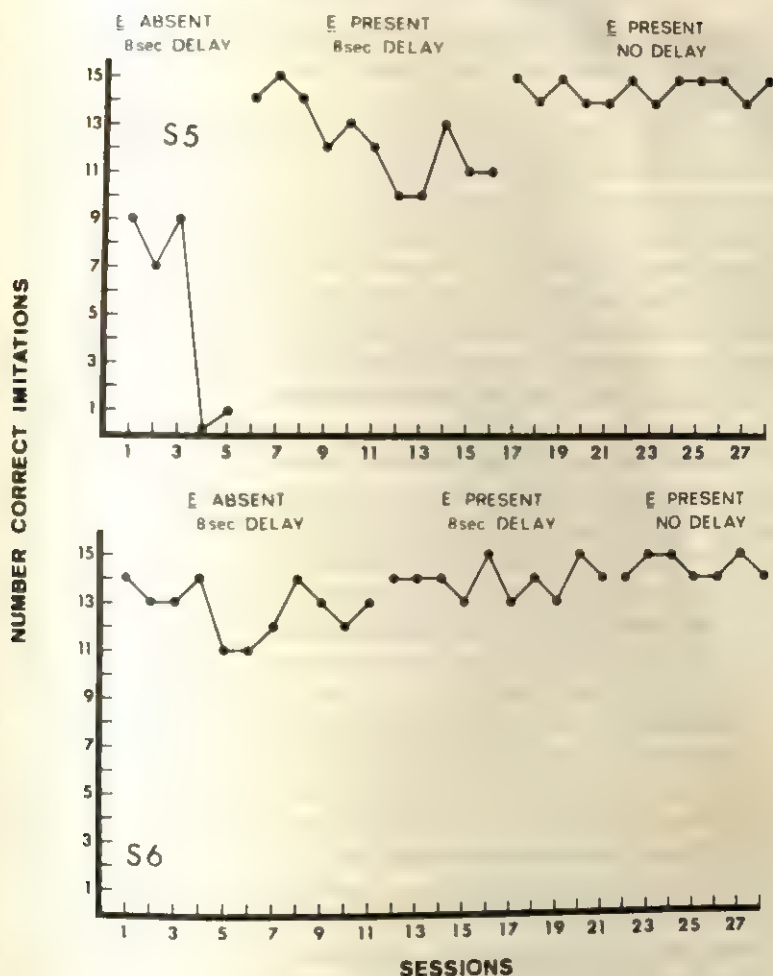


FIGURE 3

NUMBER OF NONREINFORCED IMITATIONS UNDER EACH CONDITION FOR SUBJECTS 5 AND 6

Anecdotal observations indicated that during the 8 sec delay, most subjects engaged in a variety of behaviors, such as partially rehearsing the modeled behavior, employing the experimental room, and being engaged in play and stereotyped activities. Some of these factors may also have contributed to the rapid decrease in imitative performance in subjects 3 and 4 after the initial session(s) of the "Experimenter Present—8 sec delay" condition. In addition, the reversal data of subject 3 suggest that his performance during



the first exposure to the conditions "Experimenter Present—8 sec delay" and "Experimenter Present—no delay" would have been different if these conditions had lasted longer.

Data obtained from the present study furthermore indicate that for most subjects, the absence of the experimenter resulted in the greatest reduction of imitative performance across conditions. Present findings also suggest that, for the population used in this study, the decrease of nonreinforced imitations was controlled less by the absence of the experimenter alone than might be inferred from the findings of Peterson and Whitehurst (15). Present data seem to indicate that the extent to which the absence of the experimenter affected imitative performance was a function of a decrement in performance due to response delay and its related components. It should be noted that subjects 1 and 6, who showed no decrease or only a slight decrease in imitative performance under "Experimenter Present—8 sec delay" condition, showed little or no reduction in imitative performance under the "Experimenter Absent—8 sec delay" condition. In contrast, subjects 2, 3, 4, and 5, whose performance was markedly reduced during the "Experimenter Present—8 sec delay" condition, also showed a great reduction of imitative performance under the "Experimenter Absent—8 sec delay" condition. In addition to response delay, the reduction of imitative behavior under the "Experimenter Absent" condition may have been caused by the experimenter's activities associated with leaving the room, such as walking out and closing the door.

It is interesting to note that two of three subjects (subjects 1 and 6) who had been involved in previous studies with the same experimenter showed a high and stable rate of imitative performance across conditions. It is not unlikely that the pre-imitation history with the experimenter (3, 8) was of influence upon these subjects' imitative performance.

In summary, it would appear that the observed decrement in nonreinforced imitative behaviors was controlled by response delay, the absence of the experimenter, and other coincidental factors (6), such as the stimuli associated with the experimenter leaving the room, the subject's activities during the delay period, and the pre-experimental history with the experimenter. Such findings tend to indicate a broader interpretation of Peterson and Whitehurst's findings (15) than just the presence or absence of the experimenter.

#### REFERENCES

1. BAER, D. M., & SHERMAN, J. A. Reinforcement control of generalized imitation in young children. *J. Exper. Child Psychol.*, 1964, 1, 37-49.

2. BAER, D. M., PETERSON, R. F., & SHERMAN, J. A. The development of imitation by reinforcing similarity to a model. *J. Exper. Anal. Behav.*, 1967, **10**, 405-416.
3. BANDURA, A. Social-learning theory of identificatory processes. In D. A. Joslin & D. D. Glass (Eds.), *Handbook of Socialization Theory and Research*. Chicago: Rand McNally, 1968.
4. BRIGHAM, T. A., & SHERMAN, J. A. An experimental analysis of verbal imitation in preschool children. *J. Appl. Behav. Anal.*, 1968, **1**, 151-158.
5. BUFFORD, R. K. Discrimination and instructional variables in the maintenance of nonreinforced verbal imitation. Unpublished Master's thesis, University of Illinois, Urbana, 1970.
6. BURGESS, R. L., BURGESS, J. M., & ESVELDT, K. C. An analysis of generalized imitation. *J. Appl. Behav. Anal.*, 1970, **3**, 39-46.
7. ELLIS, N. R. The stimulus trace and behavioral inadequacy. In N. R. Ellis (Ed.), *Handbook of Mental Deficiency*. New York: McGraw-Hill, 1963.
8. FLANDERS, J. P. A review of research on imitative behavior. *Psychol. Bull.*, 1968, **69**, 316-337.
9. JOHNSON, G. O., & BLAKE, K. A. Learning performance of retarded and normal children. Syracuse University Special Education Rehabilitation Monograph No. 5, Syracuse, New York, 1960.
10. KANTOR, J. R. *Interbehavioral Psychology*. Bloomington, Ind.: Principia Press, 1958.
11. LITROWNIK, A. J. Observational learning in retarded and normal children as a function of delay between observation and opportunity to perform. *J. Exper. Child Psychol.*, 1972, **48**, 117-125.
12. LOVAAS, O. I., BERBERICH, J. P., PERLOFF, B. F., & SCHAEFFER, B. Acquisition of imitative speech by schizophrenic children. *Science*, 1966, **151**, 705-707.
13. METZ, J. R. Conditioning generalized imitation in autistic children. *J. Exper. Child Psychol.*, 1965, **2**, 389-399.
14. PETERSON, R. F. Some experiments on the organization of a class of imitative behaviors. *J. Appl. Behav. Anal.*, 1968, **1**, 225-235.
15. PETERSON, R. F., & WHITEHURST, G. J. A variable influencing the performance of generalized imitative behaviors. *J. Appl. Behav. Anal.*, 1971, **4**, 1-9.
16. RISLEY, T. R., & REYNOLDS, N. J. Emphasis as a prompt for verbal imitation. *J. Appl. Behav. Anal.*, 1970, **3**, 185-190.
17. STEINMAN, W. M. The social control of generalized imitation. *J. Appl. Behav. Anal.*, 1970, **3**, 159-167.
18. ———. Generalized imitation and the discrimination hypothesis. *J. Exper. Child Psychol.*, 1970, **10**, 79-99.
19. STEINMAN, W. M., & BOYCE, K. D. Generalized imitation as a function of discrimination difficulty and choice. *J. Exper. Child Psychol.*, 1971, **11**, 251-265.
20. THOMPSON, C. W., & MARGARET, A. Differential test responses of normals and mental defectives. *J. Abn. & Soc. Psychol.*, 1947, **42**, 284-293.
21. WAXLER, C. Z., & YARROW, M. R. Factors influencing imitative learning in preschool children. *J. Exper. Child Psychol.*, 1970, **9**, 115-130.

California School of the Deaf  
3044 Horace Street  
Riverside, California 92506



## AN INVESTIGATION OF THE RELATIONSHIP BETWEEN STIMULUS EXPLICITNESS AND ENTERING BEHAVIOR IN FACILITATING STUDENT ACHIEVEMENT\*

*University Division of Instructional Services, The Pennsylvania State University*

---

THOMAS C. ARNOLD AND FRANCIS M. DWYER<sup>1</sup>

---

### SUMMARY

The purpose of this study was to investigate the relative effectiveness of specific media attributes (visual) on S performance on criterion tests measuring different levels of understanding. One hundred seventy-one Ss were assigned to one of three entering behavior groups (high, middle, and low) as a result of their performance on a pretest designed for this purpose. Ss participated in their respective presentations and received two criterion tests. Analyses indicated that (a) there was a significant relationship between entering behavior and performance on the criterion tests; (b) no significant relationship was found to exist between the level of stimulus explicitness and achievement on the criterion tests; and (c) no significant interactions were found to exist between entering behavior and type of visualization.

---

### A. INTRODUCTION

Considerable attention is currently being focused on the processes required to implement today's curriculum goals in the most efficient manner possible. One problem inherent in many curriculums is how to teach facts, concepts, and principles that require Ss cognitively to manipulate abstract representations of reality. A solution commonly employed to solve this problem has been to illustrate visually the content being presented. However, considerable research (1, 4, 6) indicates that much of this visualization is prepared with little regard for individual differences among Ss regarding the type of visualization that would be most effective in facilitating achievement of different educational objectives. Thus, the purpose of this study was (a) to investigate the relative effectiveness with which two types of visual illustrations facilitated Ss' achievement on three criterion tests measuring different

---

\* Received in the Editorial Office on March 12, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Reprints are available from the second author at the address shown at the end of this article.

educational objectives; and (b) to measure the instructional effect of different types of visuals on the achievement of Ss differing in entering behavior (that is, prior learning experiences).

## B. METHOD

### 1. *Treatment Groups*

One hundred seventy-one Ss at The Pennsylvania State University were assigned to one of three entering behavior groups (high, middle, and low) as a result of their performance on a pretest designed for this purpose. In order to insure for significant difference between entering behavior groups and to provide for reliability in assignment to the different groups, only those Ss identified as possessing high and low entering behaviors were used in this study. The assignment to these groups was performed by analyzing the cutoff scores in a procedure recommended by Helmstadter (3) and Glass and Stanley (2) involving the establishment of confidence limits about a student's obtained score. The procedure resulted in the probability of .95 that the Ss used in this study were in fact assigned to the correct entering behavior group. Ss in each of the behavior groups received their instruction in a textbook-like format consisting of 37 paragraph-like frames on 8½-inch by 5½-inch sheets; each frame contained a 2½-inch by 3½-inch illustration designed to complement the printed instruction. The instructional content selected for this study concerned the human heart, and the instructional units used were modifications of similar units employed by Dwyer (1) in a previous study.

Ss in each of the two entering behavior groups were then randomly assigned to one of two treatment groups. These treatment groups received identical written presentations; however, the two groups received visual illustrations containing different degrees of stimulus explicitness. (In this study, stimulus explicitness was defined as the amount of information contained in the visual.) Ss in Treatment I received simple line drawings, and Ss in Treatment II received detailed, shaded drawings. The line drawings contained a low level of stimulus explicitness, and the detailed, shaded drawings contained a higher level of stimulus explicitness.

### 2. *Procedure and Criterion Measures*

Each S in each treatment group received as a pretest a 20-item multiple choice identification test developed by Dwyer (1) and designed to assess S's achievement at the knowledge level of cognitive ability. Ss then interacted



with their respective instructional presentations and received two individual criterion tests. Test items on the individual criterion tests were combined into a 44-item total criterion test. The objectives of the criterion tests were as follows: (a) test of knowledge—to measure learning that requires knowledge of particular referents as well as facts that can be isolated as separate and discrete elements; (b) test of comprehension—to measure the S's ability to understand the complex relationships and processes that exist among the individual parts of the heart; (c) total test of understanding—to measure the degree to which objectives dealing with the recall of knowledge and the development of intellectual abilities were achieved by Ss. The Kuder-Richardson Formula 20 reliability coefficients for the three criterion measures were as follows: (a) knowledge .80; (b) comprehension .65; and (c) total criterion test .85.

### C. DESIGN

The data collected from this investigation were analyzed by means of a  $2 \times 2$  factorial design. The two factors investigated were entering behavior and visual explicitness, with each factor being divided into two subfactors. That is, the relationship between two levels of entering behavior and two levels of stimulus explicitness in visuals was investigated to ascertain their effectiveness on achievement of different educational objectives. In addition, comparisons among the individual means, via the two-way analysis of variance procedure, were conducted to determine if there existed interaction between entering behaviors and levels of stimulus explicitness.

### D. RESULTS

Three conclusions were derived from the data obtained in this study:

1. There was a significant relationship (at the .01 level) between entering behavior and performance on criterion tests. Those students whose prior experience and knowledge of the human heart were high performed more effectively than those with low entering behavior regardless of the type of visual illustration they received.
2. No significant relationship (at the .05 level) was found to exist between the level of stimulus explicitness and achievement on the criterion tests. This was interpreted as meaning that visuals possessing either of the stimulus explicitness levels were equally effective in improving achievement of identical objectives for each of the entering behavior groups.
3. No significant interactions were found to exist between entering behavior and type of visualization.

## E. DISCUSSION AND SIGNIFICANCE

This study is significant because of the new trends in education utilizing individualized instructional techniques complemented via various types of visualized instructional materials. One current theory (5) suggests the existence of a relationship between *S*'s entering behavior and the instructional explicitness of visual illustrations used to facilitate *S* achievement of specific educational objectives. This theory implies that *S* achievement on an instructional task will be dependent on an individual's entering behavior and the explicitness of the content material to be presented.

Within the limitations under which this investigation was conducted, the following implication may be derived: *S*s possessing high entering behaviors achieve highest on criterion tests (and *S*s with low entering behaviors achieve lowest on criterion tests) regardless of the type of visualized presentation they receive. Consequently, educators may not need to design and produce different types of visualized materials for *S*s possessing differing entering behaviors. These results contest the use of Salomon's Theory of Stimulation (5) as a viable guide for educators to consider in the production of visual aids to be used for instructional purposes.

## REFERENCES

1. DWYER, F. M. A Guide for Improving Visualized Instruction. State College, Pa.: Learning Services, 1972.
2. GLASS, V., & STANLEY, J. C. Statistical Methods in Education and Psychology. Englewood Cliffs, N.J.: Prentice-Hall, 1970.
3. HELMSTADTER, G. C. Principles of Psychological Measurement. New York: Appleton-Century-Crofts, 1964.
4. LESSER, S., & SCHUELER, H. New media research in teacher education. *Audiovis. Commun. Rev.*, 1966, 14, 318-361.
5. SALOMON, G. What does it do to Johnny? A cognitive-functionalistic view of research on media. In G. Salomon & R. E. Snow (Eds.), *ViewPoints*. Bulletin of the School of Education, Indiana University, Bloomington, Spring, 1970.
6. SNOW, E., & SALOMON, G. Aptitudes and instructional media. *Audiovis. Commun. Rev.*, 1968, 18, 341-357.

*University Division of Instructional Services  
The Pennsylvania State University  
University Park, Pennsylvania 16802*

## AN EXPLORATORY STUDY OF INDIVIDUAL DIFFERENCES IN PERCEPTUAL CENTERING AND DECENTERING\*

*University of Wisconsin, Madison*

---

ARIE COHEN AND FRANK H. FARLEY<sup>1</sup>

---

### SUMMARY

Perceptual centering and decentering, or internal *versus* external locus of perception, was studied as a function of individual differences in cognitive style, locus of control, and approval motivation (defensiveness?), with adolescent male Ss. Locus of perception was operationalized as the S's responses to discrete tactile stimulation using a new technique to infer an internal *versus* external orientation. Locus of perception as measured was found to be a significant function of all three individual difference variables [multiple correlation  $R = .70$  ( $p < .05$ )]. Implications for theoretical analyses of processes of centering and decentering were suggested.

---

### A. INTRODUCTION

An S's preference for the E's frame of reference or his own frame of reference in interpreting the orientation of laterally asymmetrical figures attached to S's body has received little attention in psychological research. Krech and Cruchfield (6) raised the possibility that Ss who interpret the symbol  $\epsilon$  written on their forehead as the letter  $\epsilon$  demonstrate greater objectivity than do Ss who interpret that symbol as the figure 3. Podell (7) proposed the notion that interpretation of such stimuli as the figure |O from an internal locus of perception, when the symbol is perceived as O|, might indicate egocentricity of the perceiver.

The purpose of the present study was to investigate some putatively relevant personality variables that might be expected to correlate with the mode of perception of such stimuli. The personality variables chosen were locus of control (8), psychological differentiation (10), and social desirability (2). Internal *versus* external locus of control was chosen as being a personality variable known to be significantly involved in an individual's orientation to

---

\* Received in the Editorial Office on March 12, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Direct requests for reprints to the second author at the address shown at the end of this article.

self and others (8) and might be expected to be related to locus of perception as presently defined. Psychological differentiation has been defined in terms of *S*'s ability to differentiate stimuli in complex stimulus contexts, with *Ss* being characterized (9) as field independent (high ability to differentiate) or field dependent (inability or low ability to differentiate). The locus of perception situation presumably requires the *S* to differentiate among stimuli and stimulus locations and would accordingly be expected to bear some relationship to such an individual difference dimension. Finally, social desirability responding, or approval seeking and approval motivation in Crowne and Marlowe's (3) terminology, would be expected to bear some relation to locus of perception. Two interpretations of the direction of such a relationship might be advanced. One would hold that high approval-seeking individuals would be more likely to adopt the *E*'s perceptual orientation than would low approval-seeking persons. Another view would lead to an opposite expectation. Approval seeking may be interpreted as defensiveness (3), and the high approval-seeker may be defined as someone preoccupied with self and its integrity. Then in the conditions of the present locus of perception paradigm, where the demands of the experiment in terms of the *E*'s approval or reproach are ambiguous, the defensive *S*'s concern with self might lead to an internal rather than external locus of perception.

## B. METHOD

### 1. Subjects

Seventeen Israeli technical school male students (age range 16-17 years) served voluntarily as *Ss*.

### 2. Procedure

Each *S* was invited to the experimental room and received the following instructions:

"Imagine to yourself that your body is a clock; your navel is the center of the clock, 12 o'clock is up, and 6 o'clock is down. I'm going to touch your body in various points while your eyes are closed, and you tell me in what hour did I touch. There is no right or wrong answer, so answer immediately the way you sense the hour."

After the instructions were read, the *S* was instructed to close his eyes, and four stimuli were then administered to the following points: the right shoulder, left leg, right hip, and left shoulder. For each *S* the number of stimuli perceived from internal locus of perception was recorded (internal locus of perception meant perceiving a stimulus on the right shoulder as



1 o'clock). Special precautions were taken to prevent Ss from discussing the experiment with one another until its completion. After all the Ss were tested for locus of perception, they were administered in a group setting the following measures: Embedded Figures Test (5), the Hebrew translation of Rotter's (8) Locus of Control Scale (4), and the Hebrew translation of Crowne and Marlowe's (2) Social Desirability Scale (1).

### C. RESULTS AND DISCUSSION

The three personality variables and the locus of perception score (the number of responses to the four stimuli from an internal locus of perception) were intercorrelated with use of Pearson product-moment correlation. The correlations between internal locus of perception and internal locus of control, Embedded Figures Test (a high score indicating field independence), and social desirability score (a high score indicating marked social desirability responding) were  $-.51$  ( $p < .05$ ),  $-.48$  ( $p < .05$ ), and  $.53$  ( $p < .01$ ), respectively. The multiple correlation between social desirability, locus of control, and locus of perception was  $.69$  ( $p < .05$ ); the multiple correlation between Embedded Figures Test, locus of control, and locus of perception was  $.58$  (NS); the multiple correlation between social desirability, Embedded Figures Test, and locus of perception was  $.62$  ( $p < .05$ ); and the multiple correlation between Embedded Figures Test, locus of control, social desirability, and locus of perception was  $.70$  ( $p < .05$ ).

The correlations between the three tests and the number of stimuli perceived from internal locus of perception suggest that the S's preference for his own frame of reference over that of *E* in interpreting such stimuli is a significant function of the S's personality.

The positive relationship between field independence and the preference for the *E*'s frame of reference might be related to the ability of Ss with high psychological differentiation to prefer an objective standpoint that is different from their own subjective frame of reference. In the same way it might be explained that individuals with internal locus of perception are those who have a passive outlook on the world and are uninterested, or unable, to invest the effort needed to perceive their body from a different frame of reference, and therefore score low on the locus of control test. The present findings, although admittedly on a small sample and restricted to adolescent males, are intriguing ones in the long-standing discussion of personality and perception, and may be extended to conceptualizations of self and other. That is, locus of perception, as presently defined, has been demonstrated to be in part a significant function of personality-cognitive



style dimensions of field dependence and generalized expectancies for internal *versus* external control of reinforcements. These results reinforce the generalization that self ("centered") *versus* other ("decentered") orientation to the world is an important characterization of the individual, being expressed in measures as presumably diverse in method as those reported here. The findings on social desirability or approval seeking are of considerable interest. They tend to be in line with the defensiveness interpretation discussed earlier, which assumed that the perceived demands of the experimental task were sufficiently ambiguous, or perhaps unimportant to an S's score, that the highly defensive S's preoccupation with self would lead to an internal locus of perception. This result might well be pursued in further research, with perhaps elaborations on the present paradigm, in order to delineate its implications for approval-motivation theory and analyses of centering-decentering.

## REFERENCES

1. COHEN, A. The Hebrew translation of Crowne and Marlowe's social desirability scale. Unpublished manuscript, Bar-Ilan University, Ramat-Gan, Israel, 1968.
2. CROWNE, D. P., & MARLOWE, D. A new scale of social desirability independent of psychopathology. *J. Consult. Psychol.*, 1960, **24**, 349-354.
3. ———. The Approval Motive. New York: Wiley, 1964.
4. FRANKEL, J. J. The Hebrew translation of the locus of control test. Unpublished manuscript, Bar-Ilan University, Ramat-Gan, Israel, 1968.
5. JACKSON, D. N., MESSICK, K. S., & MAYERS, C. T. Evaluation of group and individual forms of embedded-figure measures of field independence. *Educ. Psychol. Meas.*, 1964, **24**, 177-191.
6. KRECH, D., & CRUCHFIELD, R. S. Elements of Psychology. New York: Knopf, 1958. P. 205.
7. PODELL, J. E. Ontogeny of the locus and orientation of the perceiver. *Child Devel.*, 1966, **37**, 993-997.
8. ROTTER, J. B. Generalized expectancies for internal versus external control of reinforcement. *Psychol. Monog.*, 1966, **8**, 1-28.
9. WITKIN, H. A., LEWIS, H. B., MACHOVER, K., MEISSNER, P. B., & WAPNER, S. Personality Through Perception. New York: Harper, 1954.
10. ———. Psychological Differentiation. New York: Wiley, 1962.

University of Wisconsin

Research and Development Center for Cognitive Learning

1025 West Johnson Street

Madison, Wisconsin 53706

## CREATIVE THINKING ABILITY, SCHOOL READINESS, AND INTELLIGENCE IN FIRST GRADE CHILDREN\*<sup>1</sup>

*Department of Educational Psychology, State University of New York at Albany; and School of Education, Indiana University*

---

NICHOLAS C. ALIOTTI AND WILLIAM E. BLANTON

---

### SUMMARY

To investigate the relationships among figural and verbal creativity, language and nonlanguage intelligence, and school readiness, a test battery of the Torrance Tests of Creative Thinking, California Test of Mental Maturity, Metropolitan Readiness Tests, and the Picture Interpretation Test was administered to 83 first grade children. Varimax rotations of the correlation matrix resulted in a three factor solution identified by relatively clear and high loadings for figural creativity, verbal creativity, and intelligence-school readiness. In contrast to previous studies, the creativity measures remained as separate and distinct factors which were relatively independent of the intelligence and school readiness measures.

---

### A. INTRODUCTION

The present study explored the relationships among figural and verbal creative thinking abilities, language and nonlanguage intelligence, and school readiness in a sample of first grade children. Specifically, two questions were asked: (a) What common variance is shared between language and nonlanguage intelligence, and figural and verbal creative thinking abilities? (b) To what extent are these measures related to school readiness?

While these questions are frequently considered in creativity research, relatively few factor analytic studies have been concerned with these questions at the first grade level. For example, studies at other grade levels have found low to moderate correlations between measures of intelligence and achievement and measures of creative thinking ability (4, 5, 6, 14).

Investigations of creative behavior in young children still remain as a relatively unexplored research area. Methodological difficulties, for example,

---

\* Received in the Editorial Office on March 12, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Based on a paper presented at the American Educational Research Association, Los Angeles, California, February 8, 1969.

have centered about the creation of test materials appropriate for use with these age levels. Furthermore, the criterion problem for creative behavior becomes even more perplexing when young children are considered. Starkweather (17), for example, has questioned whether the criteria used for older children and adults can actually be differentiated among young children. Unlike Torrance and others who have used a statistical infrequency criterion (8, 18, 23), Starkweather uses each child as his own measure of originality; that is, the most original child is the one who produces the greatest number of varied responses.

In general, early attempts to research creative behavior in children have consisted of subjective observations of children's play or work. Creative activities of preschool children have been described and, in some cases, analyzed with respect to their drawings and paintings (2, 10, 11, 13), storytelling (15), and block-building and clay modeling (3, 13). More recently, the development of paper-and-pencil instruments (17, 18, 23) has again stimulated creativity research with young children.

## B. METHOD

The subjects of this study were 83 white first grade children (46 males and 37 females) enrolled in five classrooms in the Lillie E. Suder School in Clayton County, Georgia. After five weeks of school, Figural Form A of the Torrance Tests of Creative Thinking was administered in classroom groups, and Verbal Form A of this battery was administered orally and individually by a research team under the supervision of Dr. E. Paul Torrance. In addition, all children were administered the Metropolitan Readiness Tests (Form A), the California Test of Mental Maturity (revised, 1963), and the Picture Interpretation Test (22). Information concerning the first four tests is available in numerous sources.

The Picture Interpretation Test is currently under development and was included in the testing battery as part of a construct validity study. The Picture Interpretation Test was developed to measure a child's ability to "read a picture." The child is presented with a novel stimulus picture and is asked to agree or disagree with statements concerning the content of the picture (1, 20).

## C. RESULTS

A total of 16 measures was obtained and intercorrelated. The resultant correlation matrix was subjected to a principal components factor analysis with unities in the main diagonal. Following the rule of Kaiser (12), ro-

tating all factors with eigenvalues values greater than one resulted in a four factor solution. Similarly, a plot of the eigenvalues using Cattell's Scree Test (7) also suggested four factors (16). A similar number of factors emerged in the case of males, but not in the case of females where five factors emerged. In order to achieve a simple structure, varimax transformations were performed for the total sample, as well as for males and females considered separately. In all cases examination of the rotated factor structures indicated that a three factor solution most nearly conformed to a simple structure.

### 1. Three Factor Solution

Table 1 presents the principal components and the rotated varimax factors for creative thinking abilities, school readiness, intelligence, and picture interpretation for 83 children. Inspection of the rotated (varimax) components analysis revealed the following interpretation.

Factor I seemed to represent a general index of school readiness and intelligence and included the measurement of skills related to success in classroom situations. On the Metropolitan Readiness Tests high loadings were obtained for Numbers (.81), Word Meaning (.72), Alphabet (.65), and Copying (.63). Factor loadings for Language and Nonlanguage IQ were

TABLE 1  
PRINCIPAL COMPONENTS AND ROTATED VARIMAX FACTORS FOR CREATIVE  
THINKING ABILITIES, SCHOOL READINESS, INTELLIGENCE, AND  
PICTURE INTERPRETATION FOR 83 CHILDREN

Measure	Unrotated			Rotated		
	I	II	III	I	II	III
Figural Fluency	—71		—53		85	
Figural Flexibility	—63		—65		89	
Figural Originality	—62		—63		87	
Figural Elaboration	—59				44	
Verbal Fluency		—83				88
Verbal Flexibility		—81				85
Verbal Originality		—76				78
Word Meaning	—59		33	72		
Listening	—55			60		
Matching	—49			55		
Alphabet	—68			65		
Numbers	—77		33	81		
Copying	—66			63		
Language IQ	—54		31	66		
Nonlanguage IQ	—40	39		50		
Picture Interpretation			49	33		
% of total factor variance				40.8	31.1	28.1
% of total variance				23.1	17.8	16.1

Note: Decimals and all loadings  $\leq .30$  have been omitted.



.66 and .50, respectively. Picture Interpretation (.33) also contributed to this factor.

Factor II was clearly defined by the figural measures of creative thinking ability. The factor loadings were .85, .89, .87, and .44 for Figural Fluency, Figural Flexibility, Figural Originality, and Figural Elaboration, respectively. This factor was remarkably distinct. The lower loadings of the Figural Elaboration may reflect the consistent finding that fluency and elaboration are less strongly associated, since fluency is sacrificed when figural productions are elaborated.

Factor III also emerged with rather clearly defined loadings and defined the verbal creativity measures. Factor loadings were .88, .85, and .78, respectively, for Verbal Fluency, Verbal Flexibility, and Verbal Originality.

## *2. Three Factor Solution for Males and Females*

Further analyses were conducted comparing males and females. *t* ratios computed for each measure resulted in statistical significance for only two of the 16 measures reported. Males performed significantly better than females on Verbal Fluency ( $p < .01$ ) and Verbal Originality ( $p < .05$ ). Interestingly, females tended to be superior to males on the figural creativity and school readiness measures, while males were superior to females on the verbal creativity, nonlanguage intelligence, and picture interpretation measures.

Table 2 presents the rotated varimax factors for creative thinking abilities, school readiness, intelligence, and picture interpretation for males and females. Inspection of the rotated (varimax) components analysis revealed essentially the same interpretations for males and females.

Factor I was bipolar with high loadings on figural creativity measures and moderate loadings on subtests of the reading readiness measures. For males the reading readiness measures were defined by Alphabet (.60), Matching (.44), and Copying (.33), while for females Copying (.44) and Listening (.40) defined the reading readiness measures.

Factor II, which defined a general index of school readiness and intelligence, was virtually identical for males and females. In the case of males, however, Picture Interpretation (.53) also contributed to this factor.

Factor III was also bipolar with high loadings on the verbal creativity measure and a moderate loading on Nonlanguage IQ. However, in the case of males Figural Elaboration (.36) and Language IQ (— .43) also contributed to this factor.

To summarize, inspection of the factor structures revealed verbal and



TABLE 2  
 ROTATED VARIMAX FACTORS FOR CREATIVE THINKING ABILITIES, SCHOOL READINESS,  
 INTELLIGENCE, AND PICTURE INTERPRETATION FOR MALES AND FEMALES

Measure	Males			Females		
	I	II	III	I	II	III
Figural Fluency	85			87		
Figural Flexibility	90			84		
Figural Originality	86			91		
Figural Elaboration	38	37	36	56	31	
Verbal Fluency			89			84
Verbal Flexibility			87			81
Verbal Originality			77			82
Word Meaning		78			67	
Listening		66		40	55	
Matching	44	40			60	
Alphabet	60	49			75	
Numbers		78			81	
Copying	33	53		44	64	
Language IQ		63	-43		57	
Nonlanguage IQ		35	-57		48	48
Picture Interpretation		53				
% of total factor variance	35.2	34.5	30.3	34.6	38.3	27.1
% of total variance	21.5	21.1	18.5	20.3	22.5	15.9

Note: Decimals and all loadings  $\leq .30$  have been omitted.

figural creativity as defined by the Torrance tests to be remarkably independent of each other. General intelligence, school readiness, and picture interpretation also emerged as a distinct and independent third factor. These patterns were found for both males and females. The clarity of the figural and verbal creativity factors would also seem to lend support to Torrance's (18) rationale for developing both a figural and a verbal form of the Torrance Tests of Creative Thinking.

#### D. DISCUSSION

The failure of the verbal creativity measures to contribute to the school readiness-intelligence and figural creativity factors is an incongruous finding. As stated previously, at higher grade levels low to moderate coefficients of correlation have been reported between measures of verbal creativity and measures of academic achievement. In a review of approximately 300 reports, abstracts, and journal reprints that employed his creative thinking battery, Torrance (19) reported that the median of 65 coefficients of correlation between creativity measures and standardized measures of school achievement was .28. Similarly, the relationship between intelligence and verbal creativity has been reported to be higher than the relationship found in the present study. In the same review, Torrance reported that of 88 coefficients

of correlation between measures of intelligence and verbal creativity, the median correlation was .21.

Despite the similarity of the male and female factor structures describing the tests administered, a number of observations regarding sex differences can be made. The current findings support previous research which has reported few sex differences in creativity test performance below the fourth grade level. While the superiority of the males on the verbal fluency and originality scores was an unexpected finding, data reported by Torrance and Aliotti (21) suggest that by the fifth grade level females excel males on the verbal creativity tests. Additionally, the superiority of females on figural creativity and school readiness appears to be consonant with previous research.

#### REFERENCES

1. ALIOTTI, N. C. Ability to "read a picture" in disadvantaged first grade children. *Reading Teach.*, 1970, **24**, (1), 3-6, 57.
2. ALSCHULER, R. H., & HATTWICK, L. B. S. A study of Painting and Personality of Young Children. Chicago: Chicago Press, 1947.
3. ANDREWS, E. G. The development of imagination in the pre-school child. *Univ. Iowa Stud. Charac.*, 1930, **3** (4), 1-64.
4. BISH, G. A study of the relationship of intelligence, achievement, creativity, anxiety, and confidence among intermediate grade pupils in a suburban area elementary school. Doctoral dissertation, George Washington University, Washington, D.C., 1964.
5. BOWERS, J. Explorations of creative thinking in the early school years: XIV. A preliminary factor-analytic study of the creative thinking abilities of children. Bureau of Educational Research, University of Minnesota, Minneapolis, 1960.
6. ———. A study of the relationships among measures of productive thinking, intelligence, and ninth-grade achievement. Doctoral dissertation, University of Minnesota, Minneapolis, 1967.
7. CATTELL, R. The Scree Test for the number of factors. *Multivar. Behav. Res.*, 1966, **1**, 245-266.
8. CROPLEY, A. J. Creativity and intelligence. *Brit. J. Educ. Psychol.*, 1966, **36**, 259-266.
9. EBER, H. W. Toward oblique simple structure: Maxplane. *Multivar. Behav. Res.*, 1966, **1**, 112-125.
10. GRIFFITHS, R. A study of Imagination in Early Childhood. London: Routledge, 1945.
11. GRIPPEN, V. B. A study of creative artistic imagination in children by the constant contact procedure. *Psychol. Monog.*, 1933, **45** (1), 63-81.
12. KAISER, H. F. The application of electronic computers to factor analysis. *Educ. Psychol. Meas.*, 1960, **20**, 141-151.
13. McDOWELL, M. S., & HOWE, S. R. Creative use of play materials by preschool children. *Child. Educ.*, 1941, **17**, 321-326.
14. OHNMACHT, F. W. Achievement, anxiety, and creative thinking. *Amer. Educ. Res. J.*, 1966, **3** (2), 131-138.
15. FITCHER, E. G., & PRELINGER, E. Children tell stories: An analysis of fantasy. New York: Internat. Univ. Press, 1963.

16. RENTZ, R. R. Maxplane: A computer program for oblique factor rotations—Users manual. College of Education, University of Georgia, Athens, 1968 (mimeo.).
17. STARKWEATHER, E. K. Problems in the measurement of creativity in preschool children. *J. Educ. Meas.*, 1964, 1, 109-133.
18. TORRANCE, E. P. Torrance Tests of Creative Thinking, Norms-Technical Manual (Res. Ed.). Princeton, N.J.: Personnel Press, 1966.
19. ———. Minnesota Studies of Creative Behavior: 1958-1966. Greensboro, N.C.: The Creativity Research Institute, Richardson Foundation, 1968.
20. TORRANCE, E. P., & ALIOTTI, N. C. The Picture Interpretation Test: Selected test correlates. Research and Development Center in Educational Stimulation, University of Georgia, Athens, April, 1968 (mimeo.).
21. ———. Sex differences in levels of performance and test-retest reliability on the Torrance Tests of Creative Thinking Ability. *J. Creative Behav.*, 1969, 3 (1), 52-57.
22. TORRANCE, E. P., & GROSSMAN, M. The Picture Interpretation Test. Athens, Ga.: Georgia Studies of Creative Behavior, 1967.
23. WALLACH, M. A., & KOGAN, N. Modes of Thinking in Young Children. New York: Holt, Rinehart & Winston, 1965.

<i>Department of Educational Psychology</i> <i>State University of New York</i> <i>Albany, New York 12222</i>	<i>Interdisciplinary Doctoral Program</i> <i>on Young Children</i> <i>Indiana University</i> <i>Bloomington, Indiana 47401</i>
---	---



# A CRITIQUE OF SARNOFF AND ZIMBARDO'S PSYCHOANALYTIC ALTERNATIVE TO A SOCIAL COMPARISON THEORY OF EMOTIONS\*<sup>1</sup>

*Department of Sociology, University of California, Santa Barbara*

---

PAUL L. WUEBBEN

---

## SUMMARY

Sarnoff and Zimbardo's (8) contention that a psychoanalytically derived conception of emotion has greater utility in predicting affiliative response to emotional arousal than does social comparison theory was disputed. It was argued that the experiment that Sarnoff and Zimbardo conducted to test their theory was inconclusive, since (a) success of experimental manipulations of independent variables, fear and anxiety, was not satisfactorily demonstrated; (b) data relevant to the major dependent variable, affiliative preference, were presented in a form that precludes unambiguous interpretation; (c) social comparison theory as extended to emotions is adequate to account for those experimental findings that are not methodologically or interpretively suspect.

---

## A. INTRODUCTION

Schachter's (9) studies of the relationship between fear and affiliative behavior have continued to provoke widespread interest among social psychologists. [See the relevant parts of reviews by Sampson (6) and Warren (12).] One of the most intriguing suggestions to come out of Schachter's work was that social comparison theory might be relevant for explaining his findings that highly fearful subjects as opposed to less fearful subjects preferred to await electric shock with others rather than alone. Schachter suggested that since fear of electric shock is a relatively ambiguous and novel emotion, a subject experiencing such fear would be motivated to seek information that might allow him to define and evaluate his feelings. One of the reasons a highly fearful subject might seek to be with others, then, would be to

---

\* Received in the Editorial Office on March 19, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Preparation of this paper was supported in part by a grant from the Committee on Research, University of California, Santa Barbara Division. The author wishes to thank Don Zimmerman and Bob Terry for their critical comments.



observe how the others were reacting to the emotion arousing situation. A subject in possession of information about others' reactions would, presumably, use such information to achieve greater understanding of his own feelings. Festinger's (2) theory of social comparison processes was thought to be potentially applicable not only to abilities and opinions, but also to "emotions."

The present note critically examines one of the many studies that have been stimulated by Schachter's work, the study reported by Sarnoff and Zimbardo (8) in their paper, "Anxiety, fear, and social affiliation." Even though Sarnoff and Zimbardo's article appeared some time ago, its importance in current thinking is, perhaps, indicated by its recent appearance in a major collection of readings by Steiner and Fishbein (11) and by its uncritical citation in other recent works: for example, Dabbs and Helmreich (1) and Warren (12). Further, Sarnoff (7) has himself discussed his study in a recent book. In any case, there can be no doubt that Sarnoff and Zimbardo's study represents a major challenge to current conceptualizations of the relationship between "emotions," cognitions, and social comparison processes: for example, Schachter (10).<sup>2</sup>

Sarnoff and Zimbardo oppose a psychoanalytically derived conception of emotion [see Freud (3, 4)] to the conception suggested by Schachter (9). Although they agree with Schachter that fear may lead to affiliation for social comparison purposes, Sarnoff and Zimbardo (8, p. 362) maintain that "The need for self-evaluation is not the *most* salient motive aroused in the experimental situations that Schachter and we employed. We do not view the cognitive need to structure a vague emotional state as the primary motive in these experiments; we see social comparison . . . merely as one of the several responses that are *instrumental* in reducing the tension associated with the . . . motives of fear and anxiety."<sup>3</sup>

For Sarnoff and Zimbardo (8, p. 356) "All emotions are consciously experienced epiphenomena of motives . . . Motive . . . is defined as a tension producing stimulus that provokes behavior designed to reduce the tension." Given this conception, fear and anxiety are motives that must be carefully distinguished, since "Each . . . requires performance of a different response

<sup>2</sup> Schachter has elaborated considerably his view of the relationship between emotional behavior and social-cognitive influences. Schachter's general formulation of the nature of emotional behavior—relatively undifferentiated physiological arousal plus defining cognition—is consistent with and grew out of his earlier suggestions. Thus Schachter's basic theory of emotion and the experimental work which it has stimulated are both challenged by Sarnoff and Zimbardo's (8) study.

<sup>3</sup> All italicized portions of materials quoted in this paper originally appeared in Sarnoff and Zimbardo's article.

for the maximal reduction of its tension" (8, pp. 356-357). According to Sarnoff and Zimbardo (8, p. 357), fear is aroused "whenever persons are confronted by an external object or event that is inherently dangerous and likely to produce pain." The overt responses that are likely are flight and attack. Anxiety, in contrast, "is typically aroused by stimuli which, objectively considered, are innocuous . . . Regardless of their content, motives whose arousal evokes anxiety share a common property; they are all *repressed*" (8, p. 357). When given a choice, a person experiencing anxiety will respond not by seeking affiliation with others, but by seeking isolation, so that he may attempt to reassert his inner self-control and thus avert the threat of possible expression of the repressed motive the arousal of which caused his *consciously* perceived motive of anxiety.

Sarnoff and Zimbardo (8, p. 357) conducted an experiment that was designed to test the following hypotheses: "(a) the greater the fear aroused, the more subjects should choose to be with others while they await contact with the fear arousing object . . . (b) the greater the anxiety elicited, the more subjects should choose to be alone while they await contact with the anxiety arousing object." According to Sarnoff and Zimbardo, their experimental findings supported their hypotheses and their theoretical statements.

The purposes of the present paper may be outlined as follows: (a) to examine critically Sarnoff and Zimbardo's experimental data and their handling of it, (b) to argue that Sarnoff and Zimbardo's conceptualizations are *not* supported by their data, (c) to suggest that social comparison theory as extended to emotions is adequate to account for those of Sarnoff and Zimbardo's findings that do not suffer from methodological or interpretive error.

#### B. SARNOFF AND ZIMBARDO'S MANIPULATIONS

Sarnoff and Zimbardo attempted to test their hypotheses experimentally with 72 male subjects assigned randomly to one of four experimental conditions: High and Low Fear, and High and Low Anxiety. All subjects were told that they were participating in a physiological investigation of cutaneous sensitivity of various parts of the body. Fear was manipulated in the standard manner. The Fear groups were told that they would be electrically stimulated ("shocked"—High Fear) later in the experimental session and that certain measurements of sensitivity of the hand would be recorded under these conditions. Anxiety was manipulated by "arousing a motive that was assumed to have been repressed by most of the subjects" (8, p. 359). Subjects in the High Anxiety condition were told that, later in the experimental

session, physiological measurements of the skin around their mouth would be taken while they sucked on various objects: e.g., nipples, baby bottles, pacifiers. This instruction was designed to arouse "‘oral libido,’ a desire to obtain pleasurable gratification by sucking on objects that are clearly related to infantile nursing experiences" (8, p. 359). Subjects in the Low Anxiety condition were told that later in the experimental session, they would have to place in their mouths (not "suck") a set of objects less directly related to "oral libido": e.g., pipes, whistles.

Subjects in all conditions were shown slides of another purported subject (a posed model) and, under a pretext, were asked to judge how ill-at-ease or upset the pictured person seemed. "Recalling the subject whom you just saw in the slides, how upset or ill-at-ease did he seem to you? Please assign a number anywhere from zero to 100 to indicate your feeling. Zero = unconcerned, at ease; 100 = extremely concerned and ill-at-ease" (8, p. 359). Subjects' responses to this question were considered to be projective measures of the extent to which manipulations had succeeded.

#### C. CRITICISMS OF SARNOFF AND ZIMBARDO'S STUDY

The first major criticism that can be made of Sarnoff and Zimbardo's study concerns the evidence for the effectiveness of their experimental manipulations. There is no question that Sarnoff and Zimbardo were successful in inducing different levels of *arousal* in the High Anxiety and High Fear groups as opposed to the Low Anxiety and Low Fear groups. However, the crucial question is whether Sarnoff and Zimbardo were successful in inducing two different *kinds* of motives in their subjects; i.e., did subjects in Anxiety conditions respond differently than did subjects in Fear conditions to the projective measure of arousal? Sarnoff and Zimbardo's demonstration of motivational differences must be considered tenuous at best. They argued that persons facing a fear stimulus should react in a relatively more homogeneous fashion than should persons facing anxiety arousing stimuli. "Since the significance of the (anxiety) stimulus depends upon its symbolic and generally idiosyncratic associations, one would expect that a stimulus which elicited anxiety from persons with relevant predispositions (repressed motives) would have less effect on those who had more adequately resolved the conflict over the expression of the same motives" (8, p. 356). Sarnoff and Zimbardo predicted, therefore, that if their manipulations were effective, subjects in Anxiety conditions would show more *variance* in response to the projective arousal measure than would subjects in Fear conditions.

The first comment that should be made concerns the operational character

of Sarnoff and Zimbardo's measure of the effectiveness of their manipulations. Acceptance of this indirect measure as valid requires that one accede to a chain of reasoning that is, in many of its parts, theoretically controversial, conceptually vague, and empirically undemonstrated. To accept the measure as valid, one must agree that (a) individuals "repress" motives associated with infantile nursing experiences and that (b) there is greater interpersonal variability in responses to stimuli that are not "inherently" noxious than there is in responses to "inherently" noxious stimuli. This paper is not the appropriate forum for an extended discussion of the scientific status of these assumptions. Present purposes are well enough served by explicitly observing that these speculations are the bases on which the validity of Sarnoff and Zimbardo's measure rests.

The second point that must be noted is that even if their measure of effectiveness of motive manipulation is held to be valid, Sarnoff and Zimbardo's application of this measure to their experimental data produced, by standard statistical test, only weak evidence that manipulations were successful. Thus, even though Sarnoff and Zimbardo state in their text that "The heterogeneity of responses in the High Anxiety group is, as predicted, greater than in the High Fear and the Low Arousal conditions" (8, p. 360), an examination of the footnote to their Table 1 shows the following statement: "Variance greater (in High Anxiety group) than in High Fear group,  $p < .10$ ;  $SD$  for High Anxiety = 24, for High Fear = 16" (8, p. 360). Two observations are pertinent, both of which may be stated in the form of questions to which Sarnoff and Zimbardo provide no answers. First, why was it not predicted (and found) that subjects in both High and Low Anxiety conditions would show more variance than would subjects in Fear conditions? (If the motive, "anxiety," had been successfully induced and if the means of measuring it were valid, one would reasonably expect that subjects in the Low Anxiety conditions would respond with as much variance as subjects in High Anxiety conditions.) Second, what is the logical status of a statistical difference between variances of High Fear and High Anxiety groups at the unconventionally high level,  $p < .10$ ?<sup>4</sup>

<sup>4</sup> It is commonly recognized that when one is testing hypotheses, as opposed to reporting *post hoc* analyses, the logic of statistical inference demands that the level at which findings will be considered to be statistically significant be set before the analyses are undertaken; differences that do not achieve that level of statistical significance must be considered to have been due to chance variations. Sarnoff and Zimbardo do not inform the reader which level of statistical significance they had decided their results should meet to enable them to reject the null hypothesis that "The manipulations did not work." One could safely assume that they chose the .10 level, but for one fact. In the rest of their paper, Sarnoff and Zimbardo consistently cite other levels of  $p$  (particularly the



Given the questionable validity of Sarnoff and Zimbardo's measure of manipulation effectiveness and given the ambiguous results of the statistical test they applied to scores on that measure, one must conclude that there is no satisfactory evidence that Sarnoff and Zimbardo succeeded in arousing in their subjects distinctive "fear" and "anxiety" motives.

The second major criticism of Sarnoff and Zimbardo's study concerns their interpretation of subjects' responses to an 11-item structured questionnaire that explored the reasons subjects gave for choosing to wait with others.<sup>5</sup> Sarnoff and Zimbardo (8, p. 363) state that "The marked difference in the importance of the reasons given for affiliation between the High Fear and High Anxiety groups is perhaps the most substantial evidence that the experimental manipulations have indeed led to the arousal of two quite different motives." Sarnoff and Zimbardo found that the following reasons for affiliation were significantly ( $p < .05$  in each instance) more important for the High Fear group than they were for the High Anxiety group: emotional comparison, extent of comparison, distraction, catharsis, and the physical presence of others.<sup>6</sup> The first comment which may be made about Sarnoff and Zimbardo's interpretation of these significant findings is that it is not clear how a given set of data can serve both as evidence that independent variables are operative *and* as dependent variables reflecting the influence of the independent variables the operation of which is inferred from those same data. Second, given the questionable status of Sarnoff and Zimbardo's primary demonstration of manipulation success, one must ask the question, "Do the found differences (considered as dependent variables) between Fear and Anxiety groups on the 11-item questionnaire reflect the existence of two dis-

---

conventional .05 level) in reporting their findings. It therefore seems justifiable to make the inference that Sarnoff and Zimbardo were not *testing* the hypothesis, "Did the manipulations work?" with their measure of manipulation effectiveness. But if this statistical analysis was *post hoc* in nature, then other *post hoc* statistical tests have the same logical status as does Sarnoff and Zimbardo's. Such a statistical test and the interpretation it entails will be presented below.

<sup>5</sup> "Since there were too few subjects choosing the alone condition, the analysis is limited to those wanting to affiliate" (8, p. 361).

<sup>6</sup> Sarnoff and Zimbardo (8, p. 361) give the following examples, among others, of the items (as administered in the Fear condition) that were used as measures of these variables. "1. (Emotional comparison) I am not sure whether I am reacting in the same way as the others to the prospect of getting shocked and would like to compare my reaction to theirs. 2. (Extent of comparison) I feel worried about getting shocked and would like to know to what extent the others are worried too." Sarnoff and Zimbardo do not reveal how these questions were worded for the subjects in Anxiety conditions where no threat of shock was present. If words such as "worried" were used, items may not have been strictly comparable for the Anxiety and Fear groups; i.e., in what sense would a person respond that he was "worried" that he would have to suck various ob-



tinctive motives, or is there another, more persuasive, explanation which fits the data?" It is the present contention that there is another explanation which is preferable to Sarnoff and Zimbardo's not only because of its theoretical parsimony, but also because of its firm foundation in empirical fact.

Please recall that Sarnoff and Zimbardo measured both extent of fear and extent of anxiety by asking subjects to view slides of a pictured "former subject" and to judge how upset they thought him to be. It is profitable to view scores on this measure in the manner suggested by Sarnoff and Zimbardo's Table 1: i.e., as measures of projected *arousal*. Examination of Table 1 shows that subjects in the High Fear condition had a mean arousal score of 42, while those in the High Anxiety condition had a mean arousal score of 31. Using the standard deviations and *ns* reported by Sarnoff and Zimbardo, the author computed a *t* test of the difference between these mean arousal scores ( $t = 1.62$ ,  $p < .11$ ). This *p* value conforms closely to the *p* value ( $p < .10$ ) of Sarnoff and Zimbardo's test of the difference in variance between High Fear and High Anxiety groups. On statistical grounds, therefore, it is defensible to entertain the hypothesis that subjects in High Fear conditions and subjects in High Anxiety conditions responded differently on the 11-item questionnaire *not* because different *motives* had been aroused in them, but rather because of their differential *extent* of arousal (considered as a unitary concept). Further, the present hypothesis seems, in fact, preferable to Sarnoff and Zimbardo's because of the relatively direct operational interpretation of the measure on which it is based and because of the existence of extensive, independent empirical evidence which shows that as emotional arousal increases, affiliation for purposes of social comparison, distraction, catharsis, etc., increases. See reviews by Sampson (6) and Warren (12).

The third criticism of Sarnoff and Zimbardo's study is perhaps the most telling; it is directed at the manner in which the measure of the major dependent variable, preference for social affiliation, was reported. After experimental manipulations had occurred, all subjects were told that since there would be a short waiting period before physiological measurements could be taken, they could have their choice of waiting during this period either alone or with other subjects. (No restrictions were placed on talking with other subjects during the waiting period.) "After indicating their preference for waiting alone or together with others, the subjects also indicated the intensity of this preference on an 'open-ended' scale in which zero represented a very weak preference and 100 a very strong preference. On this relatively unstructured scale there was as much as 175 points of difference be-

tween subjects (from '75-alone' to '100-together')" (8, p. 358). Thus Sarnoff and Zimbardo had independent measures of affiliative *preference* and of *intensity* of preference. But in reporting their findings with respect to affiliative preference, Sarnoff and Zimbardo used their *intensity* measure as an additional criterion on the basis of which to classify subjects. In Table 2 (8, p. 361), Sarnoff and Zimbardo oppose those who chose to wait with others to those who chose to wait alone *or* chose to wait with others but at zero intensity.<sup>7</sup> In other words, Sarnoff and Zimbardo classified together those subjects who expressed a desire to wait alone and those who *actually chose to wait with others* but did not indicate, on the intensity measure, that they were other than basically indifferent to the conditions under which they waited. At no point do Sarnoff and Zimbardo reveal the distribution of subjects who indicated that they preferred to wait alone rather than with others, nor do they indicate any theoretical or methodological reason for the rather unusual manner in which they do display their findings. One cannot tell, from Sarnoff and Zimbardo's presentation, how their major dependent variable was affected by experimental treatments.<sup>8</sup>

Sarnoff and Zimbardo also present in Table 2 scores they call "Mean Affiliation Strength"; these scores were derived from the open-ended intensity questions previously described. According to the footnote to Table 2, "The larger the (Mean Affiliation Strength) score, the greater the affiliation tendency; isolation intensity score subtracted from affiliation intensity score" (8, p. 361). The comparison of primary interest is again between the mean score of the High Anxiety group, 8.0, and the mean score of the High Fear group, 51.0.<sup>9</sup> That these two means are different from each other by a conventional statistical criterion is not in question. But Sarnoff and Zimbardo's interpretations of the significant difference is questionable. Accord-

<sup>7</sup> Nineteen High Fear subjects chose to wait with others as opposed to one who chose to wait alone or with others with zero intensity; 10 High Anxiety subjects chose to wait with others as opposed to 12 who chose to wait alone or with others at zero intensity.

<sup>8</sup> It is instructive to observe that Sarnoff and Zimbardo did not group together subjects who chose to wait alone with zero intensity and subjects who chose to wait with others. This would have been equally "legitimate" as the way in which Sarnoff and Zimbardo classified subjects, although it may not have produced a distribution as favorable to the hypothesis. It should be mentioned, further, that it would not be persuasive for Sarnoff and Zimbardo to argue, as they might, that subjects who choose together but with zero intensity are essentially saying "I don't care," and should, therefore, be classified with those who want to be alone. Gerard and Rabbie (5) have effectively argued that the response, "I don't care," to the question, "Do you want to wait alone or with others?" is *not* psychologically equivalent to a positive preference for waiting alone.

<sup>9</sup> Sarnoff and Zimbardo (8, p. 361) did not find High and Low Anxiety groups to differ on this measure. "While the mean difference between High and Low Anxiety was even larger than that between the Fear conditions, it only approached significance ( $p = .16$ ,  $t = 1.46$ ) due to the marked heterogeneity of variance of the High Anxiety groups."

ing to Sarnoff and Zimbardo, this finding supports their contention that highly anxious subjects prefer isolation to affiliation. And indeed, their theoretical discussion seems to require such an outcome. However, a more accurate description of the finding is that subjects in High Anxiety conditions do not prefer with great intensity either waiting alone or waiting with others, but rather are indifferent to the conditions under which they wait. (The data show a slight tendency for subjects in High Anxiety conditions to prefer being with others rather than alone.) Thus Sarnoff and Zimbardo have demonstrated not that subjects in High Anxiety conditions prefer isolation, as their theory requires, but that subjects in High Anxiety conditions do not care whether or not they are with others.

#### D. SOCIAL COMPARISON THEORY: AN ALTERNATIVE EXPLANATION

The argument, to this point, may be outlined as follows: (a) Sarnoff and Zimbardo do not present convincing evidence that their experimental manipulations succeeded. (b) Sarnoff and Zimbardo's 11-item questionnaire data are best understood not in terms of their theory, but rather in terms of an alternative hypothesis which is both parsimonious and firmly based in empirical fact. (c) Sarnoff and Zimbardo's presentation of their data does not permit unambiguous understanding of findings related to their major dependent variable, affiliative preference. (d) Findings based upon Mean Affiliation Strength scores are best described as showing that subjects in High Fear conditions preferred waiting with others, while subjects in High Anxiety conditions were essentially indifferent to whether they waited alone or with others. (e) Sarnoff and Zimbardo's theoretical statements are not congruent with their finding that subjects in High Anxiety conditions were indifferent to affiliation.

If the foregoing contentions are well taken, the question of interest is, "How can the affiliation indifference of subjects in the High Anxiety condition be understood?" It is the present contention that this finding is congruent with social comparison theory as extended to emotions. It should be stated that the hypotheses involved in this contention must, of course, be regarded as *post hoc* until they are subjected to direct empirical test. It should also be recognized, however, that the present hypotheses, to the extent that they are derivable from social comparison theory, rest on a relatively firm foundation of empirical findings.

Social comparison theory is compatible with Sarnoff and Zimbardo's affiliation strength finding in two ways. The first of these has already been mentioned in connection with the second major criticism of Sarnoff and Zim-

bardo's study. In that discussion it was observed that subjects in High Fear conditions were more aroused than were subjects in High Anxiety conditions. If it be agreed that the more aroused a person is, the more important (salient) his arousal seems to him, social comparison theory as extended to emotions would predict the found difference between the High Fear and High Anxiety subjects. The relevant statement, as applied to opinions and abilities, appears in Festinger's paper (2, p. 130) in the following form: "In general, the more important the opinion or ability is to the person, the more related to behavior, social behavior in particular, and the more immediate the behavior is, the greater will be the drive for evaluation." Schachter (9) and others—for example, Warren (12)—have shown empirically that subjects who are most "emotionally" aroused are those who show the most intense preference for waiting with others.

The second way in which social comparison theory is congruent with Sarnoff and Zimbardo's affiliation strength finding is more circuitous and at the same time more interesting. To explicate it, certain of Sarnoff and Zimbardo's misunderstandings of social comparison theory must be noted. Sarnoff and Zimbardo (8, p. 356) state, erroneously, that "(social comparison theory) requires the assumption that *all* emotive states have the same effect on affiliative behavior." According to Sarnoff and Zimbardo, social comparison theory as extended to emotions would predict that whenever an individual is experiencing a novel, ill-defined emotion, he will seek to be with others. But a careful reading of social comparison theory shows that it predicts that a subject will seek to be with others only if he believes that he can gain self-relevant information by being with them. A subject will not perceive that he can gain self-relevant information by affiliating with others who, he believes, have not undergone the same emotion arousing experience as he *or* who have responded to the experience in a different way (with a different "emotion") than he has. The relevant statement, as applied to opinions and abilities, appears in Festinger's paper (2, p. 123) in the following form: "A person will be less attracted to situations where others are very divergent from him than to situations where others are close to him for both abilities and opinions." Schachter (9, Ch. 3) has shown, empirically, that highly fearful subjects express a desire to be with others *only if* they believe that those others have *just* undergone the *same* emotion arousing experiences as the subjects. This finding has since been replicated by Zimbardo and Formica (13).

Contrary to Sarnoff and Zimbardo's contention, then, both Schachter's



(9) experimental findings and social comparison theory as extended to emotions suggest that in order to predict that an individual will react to an emotion arousing situation by choosing to wait with others, one must know not only that he is experiencing an ill-defined emotion, but also that he has the belief that others are experiencing *similar* emotions. Two related questions must thus be asked about subjects in Sarnoff and Zimbardo's High Anxiety condition: (a) Was the "emotion" that they experienced a novel one, ill-defined for them? (b) Did the subjects believe that those with whom they could wait were experiencing similar emotions? Sarnoff and Zimbardo offer no direct evidence that is relevant to the first of these questions.<sup>10</sup> To expedite the discussion, however, let us accept their assumption that subjects in the High Anxiety condition were experiencing an emotion that was at least as ill-defined for them as was the emotion the subjects in the High Fear condition were experiencing. Then the answer to question (b) must be that Sarnoff and Zimbardo are probably correct in asserting that a subject who was told that he would later be asked to suck on various representations of the female breast would react to this information by believing that the arousal he was experiencing was *somehow* "inappropriate" (8, p. 357): i.e., of a different quality from that experienced by others and thus normatively unusual.<sup>11</sup> If this is true, then it follows that each subject in Sarnoff and Zimbardo's High Anxiety condition believed that the other subjects in that

<sup>10</sup> The important question of the relative ambiguity of various social-cognitive labels for "emotion" is one that has not been systematically explored in social comparison theory as applied to emotions. In empirical work, investigators have usually simply made the assumption that a given experimental situation would, because of the presumed low probability of its occurrence outside of the laboratory, induce in subjects a physiological state that would be ambiguous in its identification and evaluation. The vacuity of this assumption has not seriously vitiated the theoretical utility of empirical findings from studies that have employed two or more variants on the same situation: e.g., studies that have consisted of situations designed to induce high and low levels of fear of electric shock. However, in Sarnoff and Zimbardo's study, the High Fear and High Anxiety experimental situations were ostensibly different. Thus the assumption that subjects in High Fear and High Anxiety conditions were likely to have the same degree of difficulty in self-definition of their physiological states must be regarded as speculative in the extreme. It is quite conceivable, for example, that subjects exposed to the manipulations contained in the High Anxiety conditions had no difficulty in deciding that their reaction to the experimenter's demand that they suck on baby bottles, etc. was one of "embarrassment." According to social comparison theory as applied to emotions, if a subject had no doubt about what he felt—e.g., if he were sure he was "embarrassed"—he would have no need for self-evaluation and hence no desire to be with others.

<sup>11</sup> It should be mentioned that in terms of social comparison theory the subject's belief that his reaction is "inappropriate" is embedded in his beliefs about how others are reacting. In Sarnoff and Zimbardo's theory the individual is less concerned with the socially deviant nature of his reaction than he is with preventing the private expression of the "repressed" motive that caused his anxiety.



condition were reacting to the situation with a different (more "appropriate") "emotion" than he was; other subjects, in other words, were not perceived as an adequate comparison group! Social comparison theory as extended to emotions would predict, given these circumstances, not that High Anxiety subjects would seek isolation, but rather that they would be essentially *indifferent* about waiting with others. It has been argued that this is the most accurate description of what Sarnoff and Zimbardo found.

### E. CONCLUSION

Sarnoff and Zimbardo (8, p. 356) stated, early in their paper, that "It seems to us that, by adopting . . . (our) assumption about the psychological properties of emotions . . . it is possible to formulate predictions concerning affiliative processes that could not have been derived from the theory of social comparison processes." It has been argued that Sarnoff and Zimbardo were wrong in this statement and that, in fact, social comparison theory as extended to emotions generates hypotheses that are more congruent with Sarnoff and Zimbardo's own experimental findings than is their theory. In presenting these hypotheses, the author agrees, finally, with one of Sarnoff and Zimbardo's (8, p. 356) first statements: "The explication of Schachter's results in terms of the theory of social comparison processes is appealingly parimonious."

### REFERENCES

1. DABBS, J. M., JR., & HELMREICH, R. L. Fear, anxiety, and affiliation following a role-played accident. *J. Soc. Psychol.*, 1972, **86**, 269-278.
2. FESTINGER, L. A theory of social comparison processes. *Hum. Relat.*, 1954, **7**, 117-140.
3. FREUD, S. *New Introductory Lectures on Psychoanalysis* (1933). London: Hogarth, 1949.
4. ———. *Inhibitions, symptoms, and anxiety* (1936). London: Hogarth, 1949.
5. GERARD, H. B., & RABBIE, J. M. Fear and social comparison. *J. Abn. & Soc. Psychol.*, 1961, **62**, 586-592.
6. SAMPSON, E. E. The study of ordinal position: Antecedents and outcomes. In B. A. Maher (Ed.), *Progress in Experimental Personality Research* (Vol. 2). New York: Academic Press, 1965. Pp. 175-228.
7. SARNOFF, I. *Testing Freudian Concepts: An Experimental Social Approach*. New York: Springer, 1971.
8. SARNOFF, I., & ZIMBARDO, P. G. Anxiety, fear, and social affiliation. *J. Abn. & Soc. Psychol.*, 1961, **62**, 356-363.
9. SCHACHTER, S. *The Psychology of Affiliation*. Stanford: Stanford Univ. Press, 1959.
10. ———. The interaction of cognitive and physiological determinants of emotional state. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 1). New York: Academic Press, 1964. Pp. 49-80.

11. STEINER, I. D., & FISHBEIN, M., *Eds.* Current Studies in Social Psychology. New York: Holt, Rinehart & Winston, 1965.
12. WARREN, J. R. Birth order and social behavior. *Psychol. Bull.*, 1966, **65**, 38-49.
13. ZIMBARDO, P., & FORMICA, R. Emotional comparison and self-esteem as determinants of affiliation. *J. Personal.*, 1963, **31**, 141-162.

*Department of Sociology*

*University of California, Santa Barbara*

*Santa Barbara, California 93106*



## FERTILITY RATIO: ITS RELATIONSHIP TO MENTAL ABILITY, SCHOOL ACHIEVEMENT, AND RACE\*

*The University of Georgia*

---

R. T. OSBORNE

---

### SUMMARY

This study related four 1970 U.S. Census variables—population, percent increase in population since 1960, percent nonwhite, and fertility ratio—to reading and arithmetic achievement and *IQ* for all 159 Georgia counties involving over 250,000 public school children in grades four, eight, and 12 in the Fall of 1971. Since the study used fertility ratio as defined by the U. S. Census (i.e., "number of children under five per 1000 women age 15 to 49"), the marital status of the parents, the legitimacy of children, and the number of childless marriages do not affect the actual fertility ratio of a county. The correlation between fertility ratio and mean *IQ* is negative for all three grades. Correlations range from  $-.43$  to  $-.54$ ; all are significant. This dysgenic trend is also seen in the *rs* between fertility ratio and reading and arithmetic achievement.

---

### A. INTRODUCTION

Waller's (12) reassuring and optimistic report of positive relationships between fertility and education, occupation, and *IQ* not only discounts the earlier studies of Burt (2), Cattell (4), and Conrad and Jones (7), but it is also discordant with the more recent works of Campbell and Kiser (3) and F. Osborn (9). Waller attributes the important difference between his findings and the earlier studies to faulty methodology. By the lifetable method, Waller calculated the intrinsic rate of natural increase and generation length for *IQ* test score and educational attainment groups, and concluded that "natural selection is favoring an increase in mean *IQ* score in the sample of the present study, or at least there is no evidence for decrease" (12, p. 135). While Waller was extremely careful to adjust his data for non-reproducing persons and generation length, his sample included no blacks, a

---

\* Received in the Editorial Office on March 19, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

significant portion of the general U.S. population believed by some to have a high birthrate and a low level of education. This limitation of Waller's study was pointed out recently by Cattell (5).

In their analysis of the 1960 U. S. Census, Campbell and Kiser (3) report the birthrate of the nonwhite population of the U. S. to be higher than that of the white population at all socioeconomic levels, except the highest, where it dropped below the white. Fertility rates are tied to age at marriage, housing, and education. The lower the level of any of these three, the higher the birthrate. F. Osborn (9) sees as dangerous, if continued for long, the trend for black families with a better education to have families smaller than corresponding whites, while low socioeconomic group blacks have more children than corresponding whites.

In the present study, U. S. Census demographic variables and school achievement and *IQ* data will be collated and analyzed to determine for Georgia the relationship between *IQ* of public school children and the birthrate (fertility ratio) of their county of residence; that is, with respect to *IQ* and school achievement, is the trend in Georgia dysgenic or eugenic? In addition to the fertility ratio, three other census variables—population, percent change in population since 1960, and percent nonwhite population—will be examined in terms of their relationship to basic educational parameters, *IQ*, and reading and arithmetic achievement.

## B. METHOD

The population of Georgia in 1970 was 4,589,575 made up of 73.9% white, 25.9% black, and .2% other races. Included in the other races were 2347 Indians, 1836 Japanese, 1584 Chinese, and 1253 Filipinos.

From the 1970 Census of Population-Georgia the following information for the 159 Georgia counties was obtained: (a) *Population*—The range was from that of a small rural county, 1924, to that of a large metropolitan county, 607,592. (b) *Percent increase since 1960 census*—The range of change was from -28.1% to +111.5%. (c) *Percent nonwhite population*—The range of nonwhite population was from 0% in one Georgia county to 73.8% in another. (d) *Fertility ratio*—The range of the fertility ratio by county was from 267 to 520.

Census figures are simple numbers or percents, except the fertility ratio which is a straightforward definition of fertility that, in one index number, accounts for nonreproducing adults, marital status of parents, etc. The U. S. Census of Population (11) defines fertility as the number of children under five per 1000 women 15 to 49 years of age.



All Georgia children in grades four, eight, and 12 (250,000) were given mental ability and reading and arithmetic achievement tests in the Fall of 1971. The results of the testing program were obtained from the published report from the State Department of Education for each Georgia county (10). Mean scores were reported by county for all three tests. For grades four and eight, verbal and nonverbal *IQs* were reported; verbal and nonverbal *T-Scores* were reported for grade 12. All achievement scores were reported in grade placements.

Four demographic variables from the 1970 census and 12 school variables (four for each of three grades) based on the results of the Statewide Testing Program were intercorrelated. Only the correlations pertinent to the fundamental question raised above will be discussed. The complete correlation matrix is available from the writer.

### C. RESULTS

During the 1960-70 decade, Georgia counties of larger populations grew at a faster rate than did the smaller counties. The increase, however, was due to in-migration, since the fertility ratios (birthrates) of the more populous counties were lower than those for smaller counties. For example, the fertility index of the largest county in Georgia was 325, while percent increase since 1960 was 9.2. For a county near the state population median the fertility ratio was 408; the percent population increase since 1960 was 5.9. For Georgia's smallest county, the fertility ratio was 393, but the 10-year increase was only 2.6%. This trend for larger counties to increase in population at a faster rate than smaller counties, in spite of the low fertility ratio, is evident throughout Georgia. The correlation between percentage gain during the decade and population is .27 (see Table 1).

Since the percentage gain in population is inversely related to minority population ( $r = .49$ ), those counties with relatively small black populations are growing at a faster rate than those with large nonwhite populations. The one county that more than doubled in population between 1960 and 1970 had only 4.7% black population. Four of Georgia's 159 counties showed a net loss of 20% or more population for the decade. The birthrate (fertility ratio) for those counties was above the state average, but the percent black population was above 55% in all four counties.

The correlation between fertility ratio and mean *IQ* is negative for both verbal and nonverbal *IQ* for all three grades; i.e., the greater the number of children under five per 1000 women from 15 to 49, the lower the mean *IQ* for the county. These *rs* are consistent for all three grades and range from

TABLE 1  
CORRELATION MATRIX FOR GEORGIA FOURTH, EIGHTH, AND TWELFTH GRADE  
SCHOOL CHILDREN FOR *IQ*, ACHIEVEMENT, AND FOUR  
1970 U.S. CENSUS VARIABLES

Variables	U. S. Census variables			
	1	2	3	4
1. Population of county	1.00	.27	— .11	— .30
2. Percent increase in population	.27	1.00	— .49	— .12
3. Percent nonwhite for county	— .11	— .49	1.00	.47
4. Fertility ratio for county	— .30	— .12	.47	1.00
5. Fourth grade reading	.23	.44	— .73	— .52
6. Fourth grade math	.22	.41	— .69	— .49
7. Fourth grade verbal <i>IQ</i>	.25	.40	— .67	— .52
8. Fourth grade nonverbal <i>IQ</i>	.27	.50	— .66	— .46
9. Eighth grade reading	.26	.54	— .82	— .54
10. Eighth grade math	.24	.47	— .80	— .52
11. Eighth grade verbal <i>IQ</i>	.31	.52	— .81	— .54
12. Eighth grade nonverbal <i>IQ</i>	.34	.53	— .80	— .51
13. Twelfth grade reading	.29	.50	— .80	— .46
14. Twelfth grade math	.33	.46	— .70	— .47
15. Twelfth grade verbal aptitude	.34	.50	— .80	— .51
16. Twelfth grade nonverbal aptitude	.36	.53	— .79	— .43

Notes:  $r = .16$ , significant at the .05 level;  $r = .21$ , significant at the .01 level.

— .43 to — .54. The  $r$ s are significant and are higher than the — .26 reported in a recent study by Cattell. It should be noted that Cattell (5) used children per family as an index of fertility instead of the census defined fertility ratio.

Since all Georgia public school children in grades four, eight, and 12 were tested and all results were used in the analysis, the uniformly high negative correlations between *IQ* and fertility cannot be the result of sampling errors.

This dysgenic trend is also seen in the  $r$ s between fertility and reading and arithmetic achievement (Table 1). The higher the county birthrate, the lower the reading and the arithmetic achievement scores of the children of the county. Again, findings are consistent for both reading and arithmetic achievement for all three grades.

#### D. DISCUSSION

With the exception of massive national testing programs, such as Project Talent (1) and the Coleman Report (6), no other attempt has been made to relate fertility, percent nonwhite population, and net gain in population to test *IQ* and school achievement on such a large scale. The demographic figures were obtained for the State of Georgia from the 1970 U. S. Census of Population, while the test scores were provided by the State Department

of Education for 250,000 public school children who were in the fourth, eighth, and twelfth grades in the Fall of 1971. The four demographic variables are reported in numbers or percents, except the fertility ratio which is defined as the number of children under five per 1000 women age 15 to 49. By use of this index number, some objections raised by other investigators are overcome. The marital status of the parents, the legitimacy of the children, and the number of childless marriages do not affect the actual fertility rate of a county or state.

In contrast with the findings in Minnesota, our results do not show a eugenic trend. All correlations between mental ability and fertility ratios are negative. All are significant. Discordance in the finding between Waller (12) and Higgins *et al.* (8) and those of Burt (2), Cattell (4), and Conrad and Jones (7) is independent of methodology, but is related to the populations sampled. Waller (12) points out in the summary of his paper that conclusions should be limited to the all white population of Minnesota. The present study is based on the 1970 population census for Georgia. Twenty-six percent of the population of Georgia is nonwhite. The discrepancies between the findings of the Georgia and Minnesota studies point up the urgent need for expert and intensive investigations of the fertility-IQ controversy as related to national programs of education, welfare, and family planning.

#### REFERENCES

1. BURKET, G. R. Selected pupil and school characteristics in relation to percentage of Negroes in school enrollment. Project Talent Office, University of Pittsburgh, Pittsburgh, Pennsylvania, 1963.
2. BURT, C. Occasional Papers on Eugenics, No. 2. London, England: Hamilton, 1946.
3. CAMPBELL, A. A., & KISER, C. V. Nonwhite fertility and family planning. In R. H. Osborne (Ed.), *Biological and Social Meaning of Race*. San Francisco: Freeman, 1971.
4. CATTELL, R. B. The fate of national intelligence: Test of a thirteen year prediction. *Eugen. Rev.*, 1951, **142**, 136-148.
5. ———. Differential fertility and natural selection in society for IQ: Some required conditions in their investigation. In press, 1972.
6. COLEMAN, J. S. Equality of Educational Opportunity. Washington, D.C.: U. S. Govt. Print. Off., 1966.
7. CONRAD, H. S., & JONES, H. E. A field study of the differential birth rate. *J. Amer. Stat. Assoc.*, 1932, **27**, 153-159.
8. HIGGINS, J. V., REED, E. W., & REED, S. C. Intelligence and family size: A paradox resolved. *Eugen. Quart.*, **9**, 84-90.
9. OSBORN, F. Races and the future of man. In R. H. Osborne (Ed.), *Biological and Social Meaning of Race*. San Francisco: Freeman, 1971.
10. STATE DEPARTMENT OF EDUCATION. Statewide Testing Program, 1971-72. Atlanta, Ga.: State Dept. Educ., 1972.
11. U. S. BUREAU OF CENSUS. Census of Population: 1970. General Population Character-

istics. Final Report PC(1)-B12, Georgia. Washington, D. C.: U. S. Govt. Print. Off. 1971.

12. WALLER, J. H. Differential reproduction: Its relation to *IQ* test score, education, and occupation. *Soc. Biol.*, 1970, **18**(2), 122-136.

*Testing and Evaluation Center*

*The University of Georgia*

*Athens, Georgia 30601*

## CHANGES IN COLLEGE STUDENT ATTITUDES TOWARD THE ARAB-ISRAEL, INDIA-PAKISTAN, AND VIETNAM CONFLICTS\*

*New York Institute of Technology and Montgomery College*

---

WILLIAM B. ROSENBAUM AND LEONARD L. ROSENBAUM<sup>1</sup>

---

### SUMMARY

College student attitudes toward United States involvement in three international conflict areas were measured at various times during the years 1965 through 1971. An examination of the mean attitude scores obtained toward United States involvement in the Arab-Israel, India-Pakistan, and Vietnam conflicts revealed some reluctance on the part of the students to have the United States involved in any of these areas of international conflict.

Younger students were most dovish concerning United States involvement in Vietnam and were less so on the issues of the India-Pakistan and the Arab-Israel conflicts, whereas older students were more dovish regarding the issues of United States involvement in the Arab-Israel and India-Pakistan conflicts and least dovish with respect to the Vietnam conflict. Attitudes toward the involvement of the United States in both the Vietnam and Arab-Israel conflicts became significantly more dovish over time.

---

### A. INTRODUCTION

The United States has been preoccupied for many years with a major international conflict situation, the war in Vietnam. This United States involvement has evoked much speculation and concern regarding the effects of that conflict on the community and its attitudes. Another source of interest and speculation has been the possible consequences of community attitudes upon the conflict and outcome of the war. Mueller (7) has, for example, analyzed public opinion poll data to investigate popular support for the wars in Korea and Vietnam. He indicates that support for the Vietnam war was high initially, but declined as a logarithmic function of American casualties. The greater opposition to the Vietnam war, he argues, reflects mainly a shift

---

\* Received in the Editorial Office on March 20, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Reprints are available from the second author at the address shown at the end of this article.



of opinion within the intellectual left. In addition, he believes that new protest techniques learned with the civil rights movement have helped the intellectual left obtain greater attention for its opposition to the Vietnamese war.

Granburg and May (2) examined the relationship of opposition to the Vietnamese war to the personality dimension of internal *versus* external control of reinforcement. They found that externals were slightly more likely to hold dovish attitudes and to have engaged in more protest actions than internals. The undergraduates they studied in 1971 were essentially dovish on the war. They found that men were more dovish than women, though the difference was not statistically significant. They further speculated, "Had the study been done in 1965, during the formative stages of the antiwar protest, different results may have been obtained, as one sensed strong feelings of efficacy among protestors at that time" (p. 158).

In a series of studies of attitudes toward contemporary issues (1, 3, 4, 13) it was found that there was reasonable homogeneity of attitudes when different generations from a noncollege, largely working-class community were sampled. In contrast, when attitudes of college students and their parents, or people of different ages in "college towns" were examined, significant differences in attitude were observed. In these studies specific conflict areas were not considered; however, an item was included in their questionnaire which was designed to measure attitudes toward war. On this question, there was no significant difference in attitude (*a*) between college and high school students (1); or (*b*) between high school students and their parents in a noncollege community (4). On the other hand, a significant difference in attitude toward war was found (*a*) between high school students and their parents in "college towns" (3); and (*b*) between college students and their parents (13). In both of these comparisons all groups had attitudes essentially on the same side of the attitude continuum. However, children were consistently even more dovish than their parents.

The American college student community has evidently been most expressive and figured most conspicuously in speculations concerning the community attitudes and international conflict situations. It is from this community that we have periodically sampled in our efforts to learn about student attitudes toward the United States involvement in other international conflict situations: i.e., the India-Pakistan conflict and the Arab-Israel conflict. We have been able to compare attitudes toward a conflict situation after a period of time has elapsed, as well as examine the extent to which such variables as age, sex, and personal involvement influence attitudes.

## B. METHOD

### 1. Subjects

In December of 1971 a total of 96 students enrolled in undergraduate psychology courses at Montgomery College in Rockville, Maryland, served as subjects for a major portion of the research presented here. These subjects were asked to complete the three attitude scales anonymously: (i.e., Arab-Israel, India-Pakistan, and Vietnam), rate the importance of these issues, and indicate their sex and their age.

Similarly, in May, 1969, a total of 240 students in Introductory Psychology classes at Montgomery College responded to the attitude scales designed to measure attitudes toward United States involvement in the India-Pakistan and Vietnam conflicts. An additional 79 students responded to the scale to measure attitudes toward United States involvement in the Arab-Israel conflict in January, 1969 (9). Finally, responses of 101 students toward United States involvement in Vietnam in April, 1965, and a comparable sample in October, 1965, were obtained (6, 10).

### 2. Attitude Questionnaires

Attitudes toward the involvement of the United States in the Arab-Israel, India-Pakistan, and Vietnam conflicts were measured by three attitude scales. Attitudes toward each of the three attitudinal dimensions were measured by asking the subjects to respond to eight seven-point Likert-type items. Scores on each attitudinal dimension could therefore range from 8, indicating a dovish posture, to 56, indicating a hawkish attitude. The eight items included in each scale were selected because of previously obtained statistically significant ( $p < .05$ ) item-total correlations (6, 9). Obtained mean item-total correlations for these scales and odd *versus* even reliability coefficients were statistically significant at the .01 level (6, 9).

### 3. Importance Ratings

Subjects were asked to indicate on nine-point Likert-type scales the degree of personal importance they attributed to resolution of each of the three conflicts. Scores could therefore range from 1 (indicating no importance) to 9 (indicating substantial importance).

## C. RESULTS AND DISCUSSION

### 1. December, 1971, Data

Attitude measures and importance ratings were obtained from the subjects and represented the dependent variables used in the following analyses. The

Lindquist (5) Type III ANOVA represents a three-way ANOVA in which one of the independent variables is a repeated measure. In this study, the different levels of Age and Sex represented different groups of subjects. The independent variable of Issue, however, represents repeated measures on the same subjects. Consequently, a Lindquist Type III ANOVA was completed for each of the two dependent measures.

*a. Attitude measures.* When the attitude measures were submitted to the ANOVA, a statistically significant ( $F = 7.34$ ;  $df = 2, 178$ ;  $p < .001$ ) interaction was obtained between Age and Issue. Evidently, the students aged 23 and younger were most dovish on United States involvement in Vietnam (24.80) and were less so on the issues of the India-Pakistan (25.27) and the Arab-Israel conflicts (26.56). In contrast, those students over 24 years old were least dovish (31.02) on United States involvement in Vietnam. These subjects were more dovish on the issues of the Arab-Israel (25.81) and India-Pakistan (25.83) conflicts.

It is interesting that this result applied regardless of sex. We might have expected older men, who were past draft age or veterans of military service, to be less dovish (31.56) than younger men (24.46) who are more vulnerable to the draft and whose personal security needs might be the compelling element in shaping their attitudes, especially toward the war in Vietnam. It is even more difficult to speculate why older women (27.86) should be less dovish than younger women (25.26) or men. It should be noted that the overall attitudes toward these three issues were rather dovish, although a significant Issue main effect was obtained ( $F = 3.49$ ;  $df = 2, 178$ ;  $p < .05$ ). As indicated, scores could range from 8 to 56. The overall mean attitude toward United States involvement in the Arab-Israel conflict was 26.17, in the India-Pakistan conflict 25.56, and in the Vietnam conflict 28.01. Regardless of age and sex, attitudes toward United States involvement in Vietnam were significantly less dovish than attitudes toward the other two conflict situations.

*b. Personal importance rating.* In an ANOVA of personal importance ratings, a significant ( $F = 7.38$ ;  $df = 1, 81$ ;  $p < .01$ ) Age main effect was obtained. Evidently, younger people, regardless of sex, reported that resolution of these international conflicts was significantly more important to them personally (5.93) than did older students (4.73).

In addition, a statistically significant ( $F = 38.95$ ;  $df = 2, 162$ ;  $p < .001$ ) Issue main effect was obtained. Regardless of age or sex, resolution of the Vietnam conflict was seen as of significantly greater personal importance (6.48) than was resolution of the other conflicts (India-Pakistan, 4.48; Arab-Israel, 4.92).

## 2. Comparisons over Time

As indicated earlier, the availability of data collected in 1969 and in 1965 provides us the opportunity to compare these recently obtained results with earlier data, so we may study the influence upon these attitudes of the passage of time. Where possible, we will consider each international conflict situation separately.

*a. Vietnam attitudes.* Attitude scores toward United States involvement in Vietnam obtained in 1969 and 1971 for each sex separately were compared in a factorial  $2 \times 2$  ANOVA. Evidently no significant changes in attitudes, or sex differences, were obtained. In contrast, it is interesting to note that in April, 1965, male students favored a firmer (49.3) military stand in Vietnam than did the females (45.4), although even the females, on the average could be described as supporting a substantially hawkish position. These sex differences were statistically significant ( $p < .01$ ). In October, 1965, when the same attitude scale was administered, male student attitudes dropped to a mean of 45.8, thus displaying a significantly less militant attitude than was obtained six months earlier, although substantially hawkish, nevertheless. It may not be irrelevant to observe that draft calls had been markedly increased in the interim. It is interesting to note, however, that all attitudes in 1965 were substantially hawkish, while those measured in December, 1971, were essentially dovish.

Personal importance ratings for both the Vietnam and India-Pakistan conflicts were obtained from the same subjects. Consequently, the more parsimonious analysis of these ratings consisted of a Lindquist Type III three-way ANOVA. The independent variables were, therefore, Sex, Time (1969, 1971), and Issue (Vietnam, India-Pakistan).

There was a statistically significant main effect for Issue ( $F = 455.69$ ;  $df = 1, 272$ ;  $p < .001$ ); i.e., the Vietnam conflict was seen as significantly more important to the respondents (7.43) than was the India-Pakistan issue (4.15).

In addition, there was a significant main effect for Time ( $F = 4.54$ ;  $df = 1, 272$ ;  $p < .05$ ); i.e., regardless of sex, subjects saw these issues as of less personal importance in 1971 than subjects in 1969 did. This finding confirms recent speculation and observations of American college campuses in which relative political tranquility and concern with academic pursuits contrast sharply with earlier political activism (8, 11, 12). Of greater interest, perhaps, is the significant interaction between Issue and Time ( $F = 30.96$ ;  $df = 1, 272$ ;  $p < .001$ ). Evidently, Vietnam was seen as less important in 1971 than it had been perceived in 1969, whereas the India-Pakistan con-



flict was viewed as more important in 1971 than in 1969. This result, however, is not surprising in light of both the United States troop withdrawals from Vietnam and the active hostilities between India and Pakistan during the Bangladesh conflict which just preceded the collection of our data.

*b. India-Pakistan attitudes.* We also measured attitudes towards United States involvement in the India-Pakistan conflict both in 1969 and in 1971. When these responses were submitted to a  $2 \times 2$  (Sex  $\times$  Time) factorial ANOVA, no significant effects were identified. Attitudes were essentially dovish for both sexes at both times.

*c. Arab-Israel attitudes.* The subjects who in January, 1969, responded to the scale to measure attitudes towards United States involvement in the Arab-Israel conflict were not identified by sex, and the impact of this variable on attitudes could not be studied. Consequently, a *t* test comparison of the mean attitude responses by subjects in 1969 (29.5) and respondents in 1971 (26.3) was made. The students in 1971 registered significantly ( $t = 2.14$ ;  $p < .05$ ) more dovish attitudes towards United States involvement in the Arab-Israel conflict than did the students measured almost three years earlier. Please note that these earlier attitudes were also essentially dovish.

In addition, students in May, 1969, and December, 1971, differed significantly in their ratings of the Arab-Israel conflict. In a  $2 \times 2$  (Sex  $\times$  Time) factorial ANOVA, the 1971 students rated the resolution of the Arab-Israel conflict as significantly ( $F = 13.86$ ;  $p < .001$ ) less important (4.95) than did the students in 1969 (6.14). Parenthetically it may be of interest to note that a major portion of the scale designed to measure student attitudes towards United States involvement in the Arab-Israel conflict consisted of items that were measures of the United States involvement in support of Israel. The students slightly disagreed with United States involvement in support of Israel at both times, and significantly more so in the December, 1971, measures. The generally anti-involvement attitudes regarding the Vietnam and India-Pakistan conflicts also seemed to apply to support for Israel.

Our data suggest that there is some reluctance on the part of the American college students studied to have the United States involved in any of these areas of international conflict, and that this reluctance has become more pronounced with the passage of time. There is a parallel decrease in the extent to which resolution of these international conflict situations is perceived as being important to our students. Further, younger students were especially dovish on the issue of United States involvement in Vietnam, whereas their older colleagues tended to be least dovish on this very issue. Evidently,



significant differences related to age and issue exist in the college student community with respect to these attitudes.

## REFERENCES

1. EMERY, A., GOSCINSKI, F., SCHROEDER, D., & LERNER, R. M. College students' attitudes toward contemporary issues. *Psychol. Rep.*, 1971, 28, 542.
2. GRANBERG, D., & MAY, W. I-E and orientations toward the Vietnam war. *J. Soc. Psychol.*, 1972, 88, 157-158.
3. LERNER, R. M., PENDORF, J., & EMERY, A. Attitudes of adolescents and adults toward contemporary issues. *Psychol. Rep.*, 1971, 28, 139-145.
4. LERNER, R. M., SCHROEDER, C., REWITZER, M., & WEINSTOCK, A. Attitudes of high school students and their parents toward contemporary issues. *Psychol. Rep.*, 1972, 31, 255-258.
5. LINDQUIST, E. F. Design and Analysis of Experiments in Psychology and Education. Boston: Houghton Mifflin, 1953.
6. MCGINNIES, E., & ROSENBAUM, L. L. A test of the selective-exposure hypothesis in persuasion. *J. of Psychol.*, 1965, 61, 237-240.
7. MUELLER, J. E. Trends in popular support for the wars in Korea and Vietnam. *Amer. Polit. Sci. Rev.*, 1971, 65(2), 358-375.
8. ROBERTS, S. V. Students now apathetic on election. *N. Y. Times*, Oct. 2, 1972, p. 30.
9. ROSENBAUM, L. L. Sex of subject and importance of issue as determinants of pretest sensitization and resistance in attitude change. (Doctoral dissertation, University of Maryland, 1970) *Diss. Abst. Internat.*, 1971 (Feb.), 31(8-A), 4267-4268.
10. ROSENBAUM, W. B., ROSENBAUM, L. L., & MCGINNIES, E. Sex differences in selective exposure? *J. Soc. Psychol.*, 1974, in press.
11. THE EVENING STAR. Homecoming queen in new Maryland University mood. Sept. 25, 1972, p. D-1.
12. U. S. NEWS & WORLD REPORT. New mood of college students. June 19, 1972.
13. WEINSTOCK, A., & LERNER, R. M. Attitudes of late adolescents and their parents towards contemporary issues. *Psychol. Rep.*, 1972, 30, 239-244.

Department of Psychology  
Montgomery College  
Rockville, Maryland 20850



## CHANGE-MAKING STRATEGIES IN CHILDREN AND ADULTS\*

*The University of Tennessee*

---

HOWARD R. POLLIO<sup>1</sup> AND ROGER T. GRAY

---

### SUMMARY

Five groups of 20 Ss each, ranging from seven years to college age, were asked to solve a series of problems involving change-making and coin equivalences. Results indicated that the number of Ss able to make change without the use of written calculations increased as a consequence of chronological age, with the greatest increase occurring between 11 and 13 years of age. In addition, it was found that although children at all levels knew the various coin equivalences, younger children used primarily dimes and pennies in making change, whereas older children used all of the coins available. This was interpreted to mean that coin equivalences are probably learned as specific stimulus-response pairs, and that in making change younger children use untransformed base 10 units, such as pennies and dimes, largely because the computational demands placed by change-making operations force them to minimize the number of noncomputational transformations used.

---

### A. INTRODUCTION

In order for a child or adult in our society to purchase goods or services, he must first learn the properties of our monetary system and then apply this knowledge to specific situations. Such learning is of theoretical as well as practical significance, for it deals with the child's mastery of a relatively simple, yet completely well specified token system. In this system all coins have an absolute value which can always be expressed in terms of the system's minimal element, the penny. Larger-valued coins, such as dimes and quarters, can be specified in more than a single way in this system; e.g., a quarter can equal 25 pennies, or two dimes and one nickel, or two dimes and five pennies, or five nickels, etc. For all coins, however, there is some minimal number of other coins that can express its value exactly, and it seems reason-

---

\* Received in the Editorial Office on March 20, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This paper was written while the senior author was in residence at The University of Guelph, Guelph, Ontario, Canada. Thanks are due Marilyn R. Pollio for her helpful reading of an earlier draft of this paper.

able to expect that children come to recognize and use this minimal number with practice in using the system.

In school, children are taught how to carry out operations on the monetary system as exercises in arithmetic rather than as exercises in the use of a well specified token system. A child's proficiency in dealing with coins, however, must come from his knowledge and understanding of the structure of the system. While a child may be taught the values and relationships among elements of the monetary system in school, he is seldom required to use this knowledge in arithmetic exercises that deal with coins *per se*. For this reason it seems that an examination of how children actually come to make change might provide some insight into how children come to understand and manipulate a closed and well specified set of token equivalences, such as in provided by the American coin system.

In the content of this task there would seem to be three primary questions that can be asked: (a) At what age does a child show that he knows the various coin equivalences? (b) At what age can he make change accurately and how does he do it? (c) At what age is he able to make change as efficiently as most adults? In order to examine these questions, Ss ranging in age from seven to 21 years were presented with a series of simple change-making problems designed to determine whether or not they could make change, the types of coins used, and the efficiency of the strategies employed.

## B. METHOD

### 1. Subjects

Eighty children from the Boys' Club of Knoxville, Tennessee, and 20 University of Tennessee male and female undergraduate students were divided into five groups according to age. The first group consisted of seven-year-olds, the second of nine-year-olds, the third of 11-year-olds, the fourth of 13-year-olds, and the fifth of college students. All seven-year-olds had completed the first grade of elementary school at the time of the experiment. Although IQs were not available, none of the subjects was either grossly retarded or grossly accelerated.

### 2. Materials and Procedure

Four stacks of coins, consisting of 100 pennies, 50 nickels, 50 dimes, and 25 quarters, were used in this experiment. All coins of the same value were placed in the same stack and separated from coins of a different value. Each S was asked to solve 10 problems: three to test his knowledge of nickel, dime, and quarter equivalences; and seven to determine how he made change.

For the three equivalence tasks Ss were shown a coin and asked to draw from any stack, other than the stack from which the original coin was taken, a coin or coins that would equal in value to the stimulus coin. After each problem, Ss were asked to replace the coins drawn. For the seven change-making problems, all Ss were told that they were to assume they were clerks in a store and that they were to make change from a dollar for a series of seven purchases costing 10 cents, 19 cents, 37 cents, 74 cents, 12 cents, 43 cents, and 56 cents, respectively. As before, Ss were instructed to replace coins in their respective stacks after each problem, and all Ss were not allowed to use written calculations. Any S who was unable to make change was told the correct answer and asked to draw that amount of money from the stacks.

### C. RESULTS AND CONCLUSIONS

#### 1. *Percentage of Ss Making Change Correctly and Type of Strategy Used*

Results showed that no seven-year-old could make change correctly, while all 13-year-olds and college students could do so. Only two of the nine-year-olds could make change, while 12 11-year-olds were able to solve the problems without being given the answer. This information is contained in the first column of Part A of Table 1. The next two columns present the number of Ss using either an addition or subtraction strategy. Those who used addition chose the smaller elements first, adding them to the stimulus value until the sum of a dollar was reached; those who used subtraction subtracted the stimulus value from a dollar and drew this amount from the stacks, choosing larger coins first, and then smaller ones as needed to complete the total amount. All nine-year-olds who could make change used subtraction, while 11- and 13-year-olds used both addition and subtraction. The percentage of college students using addition and subtraction was much the same as that of the 13-year-olds. For all age groups there was a greater percentage of Ss using subtraction than addition.

#### 2. *Use of Specific Coins*

All Ss at all age levels were able to do the equivalence problems successfully. With the exception of three of the seven-year-olds, all Ss used the smallest number of coins possible: i.e., 2 dimes and a nickel for a quarter, two nickels for a dime, etc. Of these three children, two used 25 pennies and one used five nickels.

Part B of Table 1 presents the number of Ss using specific coins as a function of age. The first column in Section B lists the number of Ss who



TABLE 1  
PROPERTIES OF CHANGE-MAKING FOR DIFFERENT AGE GROUPS

Age	Part A: Number of subjects			Part B: Specific coins used				All coins	Part C: Efficiency of solution for all Ss	
	Making change correctly	Using addition	Using subtraction	Pennies only	Pennies & dimes only	Pennies, dimes, & nickels only	Two quarters, Pennies, dimes		Mean	Range
7	0	0	0	3	14	1	2	0	.44	.52
9	2	0	2	1	11	4	2	2	.68	.93
11	12	4	8	0	2	5	3	10	.82	.41
13	20	8	12	0	0	0	0	20	1.00	0
College	20	7	13	0	0	0	0	20	1.00	0

used only pennies in solving the change-making problems. For example, any S using this method would have used 88 pennies for a problem requiring an answer of 88 cents. The second column in Part B presents the number of Ss who used only pennies and dimes in solving the various problems. If, for example, an answer to a problem required 88 cents, Ss in this category would have used eight dimes and eight pennies. The third column in Part B presents the number of Ss using only dimes, pennies, and nickels so that 88 cents involved eight dimes, one nickel, and three pennies. For problems involving an answer that could be given in dimes and fewer than five pennies, Ss in columns two and three used the same coins; pennies and dimes. The fourth column lists the percentage of Ss who used exactly two quarters, and  $n$  pennies and  $n$  dimes when adding up values greater than 50 cents. For 88 cents Ss in this group would have used two quarters, three dimes, and eight pennies. The last column includes Ss who used all coins in making change.

The seven-, nine-, and 11-year-olds used a greater variety of change-making methods than either the 13-year-olds or the college students. Most of the seven-year-olds used only pennies and dimes, while none used all of the coins. Eleven of the nine-year-olds used pennies and dimes, and only two used all coins. As Table 1 shows there is a sudden change in coin usage at age 11, with fully 50% using all coins. All of the 13-year-old and college student subjects used all coins when making change.

The results presented in Part B include data for Ss who were able to make change correctly, as well as for those who had to be told the correct answer. An examination of the data for Ss who were able to make change correctly showed that of the two nine-year-old children who could make change, one used the two quarters, pennies, and dimes pattern, while the other used all of the coins. Of the 12 11-year-olds who could make change, two used only pennies and dimes, while four used all coins. For the remaining cases, three children used quarters, pennies, and dimes, and three used pennies, dimes, and nickels. The data for those Ss who couldn't make change correctly showed that only one child in the nine-year-old group used all coins. Of the 11-year-old children who couldn't make change, six none the less used all coins. All other Ss in both groups used dimes and pennies.

### 3. *Efficiency of Solution*

Part C of Table 1 presents the mean efficiency of solutions as a function of age. These values were obtained by dividing the minimum number of coins required to solve a problem by the number of coins actually used. If a problem was solved with use of the minimum number of coins needed, S's

efficiency score was 1.00, while if a problem was solved with twice the minimum number of coins needed, *S*'s efficiency score was .50. The mean efficiency for each individual was computed, as well as the mean for each age group. As can be seen, these efficiency scores increase systematically as a function of age, with the nine-year-olds producing the greatest range of values. By age 13 all *Ss* always made change in the most efficient way possible.

The results of the present experiment show that there is a sharp increase in the child's ability to make change without the use of written calculations, between the ages of 11 and 13. While only two of the nine-year-olds were able to make change correctly, most of the 11-year-olds and all of the 13-year-olds were able to do so. As expected, the two strategies used were those of addition and subtraction. While most *Ss* used subtraction, approximately one-third did use addition. Since addition is the simpler of the two methods (i.e., involves fewer calculational steps), subtraction would seem to be preferred largely because it is the one taught in school where money problems are treated as exercises in simple arithmetic. For this reason the addition method is never really widely used, reaching a maximum of about 40% for 13 year-old-children.

Although a number of different coins could have been used in making change, seven- and nine-year-old children preferred pennies and dimes, while older *Ss* used all coins and as many large ones as possible. Younger children seemed to prefer coins that were consonant with their experience involving base<sub>10</sub> numbers. By using only pennies and dimes, younger subjects essentially were counting by ones and tens. If quarters and nickels were used, it would have been necessary for them to count by number progressions whose properties were not as well-known as those of units (pennies) and tens (dimes). While most of the younger subjects knew that a quarter equalled two dimes and one nickel, they used pennies and dimes when adding up a sum of money in answering problems. Largely because of this, older *Ss* were more efficient than younger *Ss* in making change, a result which also characterizes more intelligent as opposed to less intelligent children (1).

Although, efficiency of solution increased directly as a function of age, nearly all children at all age levels were able to provide the most efficient solutions to equivalence problems. This seems to demonstrate that while younger children knew the various coin equivalences, they were unable to make use of these in solving specific problems. This means that a child who uses a nickel in defining a quarter will not use a nickel in making change, preferring instead to use pennies and dimes. Coin equivalences seem to be

learned as specific stimulus-response pairs (i.e., "a quarter equals two dimes and a nickel"), and not readily available for use in more demanding computational tasks, such as those required in making change.

More generally, these results suggest that knowing a specific S-R equivalence in the context of one task does not automatically guarantee that this equivalence can be used in a more complicated task. Rather it seems that under the pressures of computation a child prefers to minimize the number of transformations involved (i.e., a value of 25 cents is left as 25 pennies rather than transformed into two dimes and a nickel, or more radically into one quarter) and, therefore, will use untransformed units such as are provided by pennies and dimes. Only at about the age of 11 or 12, when the child has finally mastered the use of such computational operations as subtraction and addition, will he be able to make use of these simplifying and time-saving equivalences in order to produce maximally efficient solutions. Without making any undue fuss, this is also the age, according to Piaget, at which the child enters the final stage of intellectual development, and perhaps the ability to use simplifying transformations in the context of other operations represents one moment of this growth. Only when a child can use simplifying transformations easily will he be able to accomplish the highly abstract thinking so characteristic of this final stage.

#### REFERENCES

1. KLAUSMEIER, H. J., & LOUGHLIN, L. J. Behaviors during problem solving among children of low, average, and high intelligence. *J. Educ. Psychol.*, 1961, 52, 148-152.

*Department of Psychology*  
*University of Guelph*  
*Guelph, Ontario, Canada*





## DISPARAGEMENT BY A SUBORDINATE, INGRATIATION, AND THE USE OF POWER\*<sup>1</sup>

*Department of Social Sciences, Clarkson College of Technology*

---

EUGENE M. FODOR

---

### SUMMARY

An industrial simulation experiment was conducted to determine how appointed student supervisors respond to ingratiation by a subordinate under circumstances of group stress: i.e., where a member of the work crew disparages the supervisor's competence and has a deteriorating impact on the morale of the remaining members of the group. Supervisors dispensed no more rewards to an ingratiator under conditions of group stress than in the absence of stress. Supervisors did exhibit a marginally significant tendency, however, to give greater pay increases to all compliant workers under group stress than in the control condition.

---

### A. INTRODUCTION

Organizational supervisors typically have at their disposal the power to grant raises, to recommend promotion, or to discharge subordinates. Whether use of these powers is governed principally by the contribution each subordinate makes to organizational goals or, instead, is affected primarily by considerations that are extraneous to these goals would seem to have important implications not only for the productivity of the organizational unit, but also for the morale of individuals within the unit. The present experiment was designed to examine various factors that might influence use of power. It was an industrial simulation employing student supervisors and was patterned in part on a study by Kipnis and Vanderveer (3). A situation was created whereby one of four workers repeatedly disparaged the supervisor's competence and thereby had a negative effect on the morale of the remaining workers. All workers were fictitious, and their work output was preprogrammed.

---

\* Received in the Editorial Office on March 22, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The research reported herein was supported by Grant MH-21568-01 from the National Institute of Mental Health. Grateful acknowledgment is expressed to Francis Sharp, Thomas Ingram, and John Quintana, who served as experimenters.

Two hypotheses were investigated, as follows:

*Hypothesis I.* Student supervisors give more pay increases and a more favorable performance evaluation to an ingratiating subordinate under conditions where a member of the work crew disparages the supervisor and has a deteriorating impact on the morale of the work crew than in the absence of these conditions. The experimental conditions hereby described are conceived as constituting a state of group stress.

*Hypothesis II.* Student supervisors give more pay increases and a more favorable performance evaluation to all compliant subordinates under conditions where a member of the work crew disparages the supervisor and has a deteriorating impact on the morale of the work crew than in the absence of these conditions.

These hypotheses were formulated on the basis of Korten's model (4) which proposes that authoritarian styles of leadership arise as a consequence of heightened goal structure occasioned by an increase in group stress. Whether a proclivity to reward compliance, or more especially ingratiation, can properly be regarded as an aspect of authoritarianism is perhaps a moot question. It was thought merely that proclivities of this nature varied in accordance with the Korten model, whether one chooses to view them as characteristics of authoritarian leadership or not.

## B. METHOD

### 1. Subjects

Recruited for the experiment were freshmen and sophomore men from the general psychology course. Students were given course credit for their participation and no one refused to cooperate. Ninety subjects figured in the final data analysis—30 in each of three experimental conditions.

### 2. Procedure

Upon arriving at the laboratory, each subject was told that he would serve as supervisor in an industrial simulation. He would direct the work of four high school students in another room (the students were actually nonexistent). These workers would ostensibly assemble Tinker Toy models from pictured diagrams—a different model on each of six five-minute trials. It was the function of the supervisor to promote worker productivity by any and all means at his disposal. In communicating with workers, the supervisor could say anything he wished; he could give directions, advice on construction, feedback on awards of pay increase or decrease, encouragement, or reprimands.

The supervisor could communicate with workers individually or in any combination by means of a one-way communication device which consisted of a microphone and also a switch and light signal for contacting each worker. Workers communicated with the supervisor through handwritten notes (which, of course, were preprogramed, since there were no actual workers). The reason given supervisors for not being permitted to see the workers and for not receiving voiced communications from them (rather than written notes) was that the experimenter did not want the supervisor's final evaluation of the various workers to be colored by direct visual and auditory impressions of them.

At the end of each trial the work output from each worker (expressed as total number of models completed) was brought to the supervisor from another room by an assistant, along with any notes from individual workers. The supervisor was given a card containing percentile norms for the model that workers were building during a given trial. He could thereby determine a percentile score for each worker on the basis of the number of models he completed. These scores were recorded by the supervisor on a sheet provided for that purpose.

Placed in front of the supervisor was a large sign authorizing him to award a 10 cent increase or decrease to each worker at the end of any given trial (he was told at the outset of the experiment that each worker would be paid one dollar for his participation). At the conclusion of the six trials, the supervisor was asked to average for each worker the percentile scores received over the six trials and to turn in these averages together with the production records kept on the four workers. He was thereupon led to another room and requested to evaluate the performance of each worker on four nine-point scales: (a) worker's ability, (b) worker's overall worth to the company, (c) willingness to rehire the worker for a second experiment, and (d) recommendation that the worker be promoted to supervisor in a future experiment. For purposes of analysis, all scale scores were later summated to obtain a single overall score.

There were three different conditions in the experiment, but all conditions shared certain common features. Man A was always a superior worker, assembling models at a rate which on the average placed him at about the 75th percentile. On trial 1 Man A wrote, "What fun. I feel like a kid again." The comment was designed to be a neutral one. Man B performed badly across all conditions, averaging at roughly the 25th percentile. Man C was an average worker, scoring at approximately the 50th percentile. On trial 4 he asked, "May I smoke?" Man D's output closely matched that of Man C,

although for purposes of realism it often ran above or below Man C's output for a given trial. Notes coming from Man D were designed to cast him in the role of ingratiation in the sense in which that concept has been treated by Jones (2). On trial 1 he wrote, "Count on me for help—what can I do?" On trial 3 he commented, "I was afraid you people would boss me around, but your assistant is a good guy and you seem to be too." And on trial 5, "You're doing a pretty good job."

Condition 3 of the experiment included these general conditions and no more. It was the control group.

Condition 2 had Man B writing comments of disparagement to the supervisor. On trial 1 he wrote, "These models are tough to put together and you're not very much help." On trial 2, "You're a lousy supervisor!" On trial 4, "I hope you're not planning to be supervisor when you finish college. You're not very good at it." These experimental manipulations were seen as creating a situation of stress for the supervisor.

Man B's comments of disparagement also occurred in condition 1, but here the situation could more properly be regarded as one of group stress rather than merely stress directed upon the supervisor, for it was apparent to the supervisor that Man B's attitude was having a negative effect on the morale of the remaining workers. Notes from various workers to the supervisor read: "I think I could do better but Man B bugs me"; "Please straighten up Man B"; and "Why did Man B ever volunteer for this work?"

At the end of the experiment all subjects were interviewed to determine whether they sensed the purpose of the experiment. The nature of the research was then explained to them, and they were pledged to secrecy. Eleven percent of the subjects refused to believe there were workers in the other room and were therefore discounted from data analysis.

### C. RESULTS

Since predictions were made in advance of actual collection of the data, the technique of analysis employed was that of planned comparison of means.

The question of whether ingratiation yields greater benefits under group stress than in the absence of stress involves a comparison of the rewards accorded Man D under condition 1, the group stress condition, against the rewards given Man D under condition 3, the control condition where stress was absent. Table 1 shows the mean increases in pay awarded to the three compliant workers, Man A, C, and D. It can be observed that Man D received a mean increase of 20.3 cents in condition 1 and a mean increase of



15.0 cents in condition 3. The test for this comparison yielded a figure of 2.41 as the value of  $F$ , which for 1 and 87 degrees of freedom approaches but does not attain significance.

TABLE 1  
MEAN PAY INCREASES (IN CENTS) AWARDED COMPLIANT WORKERS

Condition	Man A	Man C	Man D	Mean
1	42.3	21.0	20.3	27.9
2	36.3	22.7	20.7	26.6
3	35.7	16.3	15.0	22.3

On the performance evaluation variable, Man D did only slightly better in condition 1 than in condition 3—a mean of 24.2 as compared with a mean of 23.2. The difference did not approach significance.

Examination of Hypothesis II requires that the rewards distributed to the three compliant workers be compared across conditions 1 and 3. The mean pay increase given Man A, C, and D in condition 1 was 27.9 cents, and the corresponding figure for condition 3 was 22.3. A value of 2.82 was obtained for  $F$ , which is significant at the .10 level.

Differences on performance evaluation between conditions 1 and 3 were virtually nonexistent. Compliant workers averaged 26.6 in condition 1 and 26.3 in condition 3.

#### D. DISCUSSION

Although the hypotheses examined in the present experiment were not fully supported by the data, a trend in the predicted direction did occur on the variable of pay increase. Why a similar trend did not appear on the performance evaluation variable is not entirely clear. It must be remembered, however, that the supervisor dispensed pay increases during the main body of the experiment when various stress factors were at their height, whereas he did not complete the performance evaluation until the very end and then in a room different from the one in which he had been directing his work crew. The impact of stress factors may not have been as salient at that later time.

Condition 2 was included in the experiment as an additional control group to determine whether effects produced in condition 1, if any, could reasonably be attributed to Man B's comments to the supervisor alone, without their producing a negative effect on the morale of the remaining workers. Since the strongest trend in the data favored support of Hypothesis II on the variable of pay increase, it seems most appropriate in considering



this question to look at pay increases given compliant workers in condition 2. The mean increase accorded compliant workers in condition 2 was 26.6 cents, which is intermediate between the 27.9 cent figure for condition 1 and the 22.3 cent mean for condition 3 but clearly closer to the mean for condition 1. It may therefore be primarily stress upon the supervisor *per se* rather than a situation of group stress that accounts for the observed trend in these data.

As in a previous study (1), there was no general tendency for Man D, the ingratiator, to be favored by the supervisor over Man C, a noningratiating worker of equivalent output to Man D. Man D, across conditions, received a mean performance evaluation of 23.8. The corresponding mean for Man C was 23.2. The difference did not approach statistical significance. On the variable of pay increase, Man C actually did a little better than Man D. The mean increase for Man C was 20.4 cents, as against 18.4 cents for Man D.

#### REFERENCES

1. FODOR, E. M. Group stress, ingratiation, and the use of power. *J. Soc. Psychol.*, in press.
2. JONES, E. E. *Ingratiation*. New York: Appleton-Century-Crofts, 1964.
3. KIPNIS, D., & VANDERVEER, R. Ingratiation and the use of power. *J. Personal. & Soc. Psychol.*, 1971, 17, 280-286.
4. KORTEN, D. C. Situational determinants of leadership structure. *J. Conflict Resolut.*, 1962, 6, 222-235.

*Department of Social Sciences*  
*Clarkson College of Technology*  
*Potsdam, New York 13676*

# ABBREVIATIONS OF WORDS USED IN THE TITLES OF JOURNALS

(One-word titles are never abbreviated. The word "the" is not used, nor its equivalent in any other language. The word "of" or its equivalent in other languages is used only to discriminate what would otherwise be identical titles in different languages. The word "and" is always used, but indicated by "&" in the Roman alphabet. Only English words are indicated here, but the corresponding words in other languages should receive a corresponding abbreviation. All abbreviations and all one-word titles should be in italics.)

Abnormal .....	<i>Abn.</i>	Japanese .....	<i>Jap.</i>
Abstracts .....	<i>Abst.</i>	Journal .....	<i>J.</i>
American .....	<i>Amer.</i>	Mathematical .....	<i>Math.</i>
Anatomy .....	<i>Anat.</i>	Measurement .....	<i>Meas.</i>
Animal .....	<i>Anim.</i>	Medical .....	<i>Med.</i>
Applied .....	<i>Appl.</i>	Mental .....	<i>Ment.</i>
Archives .....	<i>Arch.</i>	Monographs .....	<i>Monog.</i>
Association .....	<i>Assoc.</i>	Neurology .....	<i>Neurol.</i>
Attitude .....	<i>Attit.</i>	Opinion .....	<i>Opin.</i>
Australian .....	<i>Aust.</i>	Orthopsychiatry .....	<i>Orthopsychiat.</i>
Behavior .....	<i>Behav.</i>	Personality .....	<i>Personal.</i>
British .....	<i>Brit.</i>	Personnel .....	<i>Person.</i>
Bulletin .....	<i>Bull.</i>	Philosophy .....	<i>Philos.</i>
Bureau .....	<i>Bur.</i>	Physics .....	<i>Phys.</i>
Canadian .....	<i>Can.</i>	Physiology .....	<i>Physiol.</i>
Character .....	<i>Charac.</i>	Proceedings .....	<i>Proc.</i>
Children .....	<i>Child.</i>	Psychiatry .....	<i>Psychiat.</i>
Chinese .....	<i>Chin.</i>	Psychoanalysis .....	<i>Psychoanal.</i>
Clinical .....	<i>Clin.</i>	Psychology .....	<i>Psychol.</i>
College .....	<i>Coll.</i>	Psychosomatic .....	<i>Psychosomat.</i>
Comparative .....	<i>Comp.</i>	Quarterly .....	<i>Quart.</i>
Consulting .....	<i>Consult.</i>	Religious .....	<i>Relig.</i>
Contributions .....	<i>Contrib.</i>	Research .....	<i>Res.</i>
Development .....	<i>Devel.</i>	Review .....	<i>Rev.</i>
Educational .....	<i>Educ.</i>	School .....	<i>Sch.</i>
Experimental .....	<i>Exper.</i>	Science .....	<i>Sci.</i>
General .....	<i>Gen.</i>	Social .....	<i>Soc.</i>
Genetic .....	<i>Genet.</i>	Statistics .....	<i>Stat.</i>
Indian .....	<i>Ind.</i>	Studies .....	<i>Stud.</i>
Industrial .....	<i>Indus.</i>	Teacher .....	<i>Teach.</i>
International .....	<i>Internat.</i>	University .....	<i>Univ.</i>
Italian .....	<i>Ital.</i>		

1. The proper sequence for the parts of your submitted manuscript is as follows: (a) text (b) references, (c) footnotes, (d) tables, (e) figures, and (f) figure legends. However, monographs start with a table of contents and may have an acknowledgment page before the text and an appendix immediately after the text.
2. Use heavy typewriter paper,  $8\frac{1}{2} \times 11$  inches, double-space *all* lines, and leave margin for editorial work. Do not use onionskin or odd sizes of paper.
3. Submit original typewritten copy and one carbon. Retain second carbon for proofing.
4. Retype any page on which written corrections have been made.
5. Do not begin a sentence with a numeral.
6. An adequate summary serving as the conclusion of the text is good form.
7. Each quotation should indicate the page number of the original source. The original publisher must give permission for lengthy quotations and use of tables or figures.
8. Do not fold your manuscript.
9. Enclose your covering letter with your manuscript. If you are unknown to the Editor, kindly state your credentials.

## FORMAT AND SPECIFIC INSTRUCTIONS

### A. TEXT DIVISIONS

- I. THE TITLES OF JOURNAL ARTICLES AND THE MAJOR SUBDIVISIONS OF MONOGRAPHS ARE PRINTED IN TEN-POINT CAPS  
CENTERED ON THE PAGE

#### A. THE NEXT SUBDIVISION TITLE IS PRINTED IN CAPS AND SMALL CAPS CENTERED ON THE PAGE

1. Then *Italics, with Principal Words, Upper and Lower Case, Centered on the Page*
- a. Then *italics, upper and lower case, 1-em run-in side head.*
- (1). Then *italics, upper and lower case, 2-em run-in side head.*
- (a). Then *italics, upper and lower case, 3-em run-in side head.*

[Further subdivision should be merged into the text without marginal indentation, and should be numbered with small letters.]

### B. REFERENCES

References should be arranged in alphabetical order by author, numbered, and referred to in the text by number (2).

The proper form of a book reference is as follows:

2. Doe, J. The Preparation of Manuscripts. New York: Holt, 1963. Pp. 400-418.

The proper form of a journal reference is as follows:

2. Doe, J. The preparation of manuscripts. *J. Gen. Psychol.*, 1963, 68, 450-462.

If in the text it is desirable to refer to a page, thus (2, p. 45).

### C. FOOTNOTES

Use as few as possible and number consecutively in the text thus.<sup>1</sup>

- <sup>1</sup> Footnote (on the separate footnote page).

### D. TABLES

Tables should always be typed on separate sheets of paper and should be Arabic numbered (Table 2). Construction should be such that vertical lines are unnecessary. Always indicate in the text the approximate insertion place for a table.

### E. FIGURES

Figures should be submitted as glossy prints of the approximate size for final reproduction (ordinarily  $4\frac{1}{2}$  inches in width). Lettering and lines should be sharp and clean. Figures should be Arabic numbered (Figure 2) in the text, with the approximate insertion point indicated. Each figure requires a legend, but all legends should be submitted double-spaced on a separate Figure Legends page.

\$39.00 per annum  
\$18.00 per volume  
Single Numbers \$9.00

BIMONTHLY  
Three volumes per year  
Immediate publication

July, 1973  
Volume 84, Second Half

Founded in 1935 by Carl Murchison

# The Journal of PSYCHOLOGY

The General Field of Psychology

POWELL MURCHISON  
Managing Editor

HELEN M. DAVIS  
Associate Editor

## Editorial Board

JOHN A. ANTOINETTI  
Life Insurance Agency  
Management Association

S. HOWARD BARTLEY  
Michigan State University

ARTHUR G. BILLS  
University of Cincinnati

E. A. BOTT  
University of Toronto

LEONARD CARMICHAEL  
National Geographic Society

IRVIN L. CHILD  
Yale University

HERBERT S. CONRAD  
U. S. Office of Education

JOHN F. DASHKILL  
Florida State University

LEONARD W. DOOB  
Yale University

If this space should be unstamped, this is the regular library edition.  
But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not  
be resold for a period of five years from date of publication.

FRANK A. GELDARD  
Princeton University

ERNEST R. HILGARD  
Stanford University

JOHN E. HORROCKS  
Ohio State University

DAVID HOTHERSALL  
Ohio State University

HEINRICH KLÜVER  
University of Chicago

NORMAN L. MUNN  
University of Adelaide,  
Australia

WARREN H. TEICHNER  
New Mexico State  
University

LEE EDWARD TRAVIS  
Fuller Theological  
Seminary

JOSEPH ZUBIN  
Psychiatric Institute,  
Columbia University

Copyright © 1973 by The Journal Press  
Provincetown, Massachusetts, U. S. A.

All rights reserved

Second-class postage paid at Provincetown, Mass.  
and additional mailing offices

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

## Back Issues

The Journal Press maintains a complete inventory of back issues for all five journals by reprinting many of them each year. Early volumes frequently consist of multiple parts, and the cost of maintaining these volumes is naturally much higher than that for two-part volumes. So our old policy of a flat price per volume cannot be continued. The following rates are in effect: single numbers, \$9.00; two-part volumes, \$18.00; three-part volumes, \$27.00; four-part and over-four-part volumes, \$36.00. The only exceptions (special issues) are starred below.

	VOLUMES	DATES	PER VOLUME
<i>Journal of Genetic Psychology</i> (Pedagogical Seminary)			
	1-4	1891-April 1897	\$27.00
	5-37	July 1897-1930	36.00
	38*	1930	18.00
	39	1931	36.00
	40 and later	1932-	18.00
<i>Genetic Psychology Monographs</i>			
	1-22	1926-1940	36.00
	23 and later	1941-	18.00
<i>Journal of General Psychology</i>			
	1-3	1923-1930	36.00
	4*	1930	18.00
	5	1931	36.00
	6 and later	1932-	18.00
<i>Journal of Social Psychology</i>			
	1-10	1930-1939	36.00
	11 and later	1940-	18.00
<i>Journal of Psychology</i>			
	1 and later	1936-	18.00

A mailing charge of \$0.30 per volume is made for foreign deliveries only. Domestic mailings are prepaid. All remittances should be by check or money order payable in U.S. funds on a United States bank. Please do not send UNESCO coupons. Advance payment is required from all individuals and from foreign purchasers. Microfilm versions and special bindings are not available.

### THE JOURNAL PRESS

2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.



273  
9/12/3

*Notice: All copyrighted material reprinted or otherwise reproduced from Journal Press publications should be given an appropriate acknowledgment. No permission is necessary for brief quotations and originally worded abstracts (less than 300-word total or 10 percent of length, whichever is least, for any one article). Written permission is required for reprinting or copying tables, figures, summaries, entire articles, or substantial portions thereof; such permission normally carries a charge of \$5 per page and is contingent upon agreement and possible stipulation of the original author or authors. Estate permission can be required. Authors requesting use of their own material are not charged. Fees may be waived for limited-copy instructional usage and for government periodicals. Special rates can be applied when justified.*

---

# The Journal of Psychology

Founded in 1935 by Carl Murchison

The Journal of Psychology was founded at the request of many psychologists throughout the world for the purpose of establishing a medium in which publication shall be immediate and in which other publication experiments may be carried out.

(1). The subscription price is \$39.00 per annum, \$18.00 per volume, \$9.00 per issue (three volumes per year).

(2). Manuscripts may be sent to any member of the Editorial Board or may be sent directly to The Journal Press. The lag in publication should not be more than a few weeks.

(3). The author will receive 200 reprints of his paper, and subscribers will receive the accumulated publications bound into half-volume units. The second half of each volume will contain the title page and table of contents for that volume. Each half-volume unit will be mailed to the subscribers at its stated date.

(4). Each author will be charged the manufacturing costs of any zincs, halftone cuts, tables, or equations. Excessive use of special fonts (references, footnotes, quotations, appendices) will also be charged.

(5). The publication charges are as follows:

1 to 4 pages	\$60.00
5 to 8 "	120.00
9 to 12 "	180.00
13 to 16 "	240.00
17 to 20 "	300.00
21 to 24 "	360.00
25 to 28 "	420.00
29 to 32 "	480.00

These charges are subject to adjustment as justified.

(6). The accumulated units will be mailed to subscribers bimonthly: January, March, May, July, September, and November. The units may vary in size to accommodate publication needs.

All back issues are available. See announcement on inside front cover.

## THE JOURNAL PRESS

2 Commercial Street

Provincetown, Massachusetts 02667

U. S. A.

# Journal of Genetic Psychology

Founded by G. Stanley Hall in 1891; reorganized by Carl Murchison in 1925

Quarterly, beginning with March. Devoted to developmental and clinical psychology, with briefly reported Replications and Refinements and occasional book reviews. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$26.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

---

# Journal of General Psychology

Founded in 1927 by E. B. Titchener and Carl Murchison

Quarterly, beginning with January. Devoted to experimental, physiological, and comparative psychology, with briefly reported Replications, Refinements, and Comments. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$26.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

---

# Journal of Social Psychology

Founded in 1929 by John Dewey and Carl Murchison

Bimonthly, beginning with February. Devoted to studies of persons in group settings and of culture and personality; special attention to cross-cultural articles and notes, to field research, and to briefly reported Replications and Refinements. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Three volumes annually. Per annum, \$39.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

---

*Each author will receive 100 unbound reprints gratis. Additional reprints may be purchased at a nominal cost when the author returns page-proof.*

THE JOURNAL PRESS  
2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.

# Club Rate for Five International Psychological Journals

	PER ANNUM
THE JOURNAL OF GENETIC PSYCHOLOGY . . . . . (Quarterly, two volumes per year)	\$26.00
GENETIC PSYCHOLOGY MONOGRAPHS . . . . . (Quarterly, two volumes per year)	\$26.00
THE JOURNAL OF GENERAL PSYCHOLOGY . . . . . (Quarterly, two volumes per year)	\$26.00
THE JOURNAL OF SOCIAL PSYCHOLOGY . . . . . (Bimonthly, three volumes per year)	\$39.00
THE JOURNAL OF PSYCHOLOGY . . . . . (Bimonthly, three volumes per year)	\$39.00
	<hr/> TOTAL \$156.00
TEACHERS AND STUDENTS EDITION, <i>Restricted Sale</i>	\$50.00

The *Teachers and Students Edition* is open to any individual psychologist, teacher, student, or physician but requires an agreement that the edition will not be offered for sale till after five years from publication and an agreement that the edition will not be substituted for an institutional subscription. Write for our Teachers and Students contract card.

THE JOURNAL PRESS  
2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.

## Genetic Psychology Monographs

Founded in 1925 by Carl Murchison

Quarterly, beginning with February. Devoted to developmental and clinical psychology. Each number one or more complete researches. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$26.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

**NOTICE.** Effective immediately we offer prior publication for certain selected monographs. Charges will be based on the announced rates for *The Journal of Psychology*. Regular monographs will carry no page charge for the first 50 pages, but for all pages above 50 there will be a charge of \$10.00 per page. Manufacturing costs of any zincs, halftone cuts, tables, or equations will be charged to authors, as will excessive use of special fonts. Two hundred reprints of each monograph will be supplied free of charge.

The listing of monographs for 1926 through 1963 will be printed in the first half of each volume of all five journals. Monographs published after 1963 will be listed in the second half of each volume.

# Genetic Psychology Monographs (continued)

VOLUME 69—January-June, 1964

- Occupational and industrial therapy in the treatment of psychiatric patients: A controlled study of efficacy in an intensive-treatment institution—M. HUNDIAK AND B. PASAMANICK  
An analysis of Carl R. Rogers' theory of personality—M. S. KRAUSK  
Personality aspects of teaching: A predictive study—H. ZIMILES, B. BIDER, W. RABINGWITZ, AND L. HAY  
Situational and personal antecedents of incipient alienation: An exploratory study—W. H. JARRETT AND A. O. HALLER
- Age changes in children's Mosaic responses from five to ten years—L. B. AMES AND F. L. ILC  
A study of the interview process: The caseworker-client relationship—H. J. MEYER, E. F. BURGATTA, AND D. FANSHIEL  
A study of introversion-extroversion as a personality dimension—R. M. ALLEN, H. M. RICHER, AND R. J. PLOTNIK  
Psychological testing with Northwest Coast Alaskan Eskimos—C. E. PRESTON  
A study in family pathology—M. B. HANS

VOLUME 70—July-December, 1964

- Structure and function—H. GUTMAN  
The Lowenfeld Mosaic Test: Norms for five-to-ten-year-old American public-school children and comparative study of three groups—L. B. AMES, F. L. ILC, AND J. AUGUST  
Social and personality correlates of children's estimates of height—J. P. SHAFFER  
The development of ethnic attitudes in New Zealand school children—G. M. VALGHAN  
The antecedents of two varieties of vocational interests—L. E. TYLER  
Personality in old age—P. E. SLATER AND H. A. SCARR  
Variables affecting experimenter bias in a group situation—R. ROSENTHAL, C. J. FRIEDMAN, C. A. JOHNSON, K. FOLE, T. SCHILL, R. C. WHITE, AND L. L. VIKAN  
Ordinal position and sex role identification—B. G. ROSENBERG AND B. SUTTON-SMITH  
Personality correlates of accuracy of self-appraisal in four-year-old children—PAUL S. WEINER

VOLUME 71—January-June, 1965

- Author index for Genetic Psychology Monographs 1926-1964—L. W. FERGUSON  
A theoretical, developmental, and clinical investigation of embarrassment—J. M. SATTLER  
School readiness as evaluated by Gesell developmental, visual, and projective tests—F. L. ILC, L. B. AMES, AND R. J. APPELL  
Rorschach scales of regression in psychosis—S. NEICKER, A. C. SLEMON, AND D. A. QLIIRK  
The effects of extra tactile stimulation on a group of institutionalized infants—L. CASLER  
Human-cyloid conditioning: The recent experimental literature—L. E. ROSS AND T. F. HARTMAN
- Social status and schoolroom aggression in third-grade children—B. TOIGO  
Overt and covert homosexual problems among male college students—L. J. BRAATEN AND C. D. DARLING  
Role conceptions of the psychiatric aid—M. P. LAWTON AND A. E. GOLDMAN  
Changes in study of values scores during the four years of college—C. W. HUNTLEY

VOLUME 72—July-December, 1965

- A study of process and method in three-year-old twins and triplets learning to read—W. FOWLER  
Individualized teaching with therapeutic aims: A methodological study—M. KOHN AND M. RUDNICK  
Ability and social adjustment at midlife of persons earlier judged mentally deficient—E. L. MILLER
- Development of incentive values in childhood—S. L. WITVOLD, D. J. TYRELL, AND L. M. LOWDEN  
Longitudinal changes in intolerance and authoritarianism for subjects differing in amount of college education over four years—W. T. PLANT  
An NDEA counseling and guidance institute: Prediction, performance, and follow-up—R. H. DANA  
The influence of language on classification: A theoretical model applied to normal, retarded, and deaf children—H. G. FURTH AND N. A. MILGRAM  
Time estimates as affected by incentive class and motivational level—M. FELIX  
Clinical psychology: An historical survey—V. S. SEXTON

VOLUME 73—January-June, 1966

- Vulnerability in infants and young children: A pilot study—C. HIEGER
- Racial preference and ego-identity among white and Bantu children in the Republic of South Africa—A. J. CREGOR AND D. A. MCPHENSON  
Mother-child interaction in cerebral palsy: Environmental and psychosocial obstacles to cognitive development—E. SHERR AND R. KASTENBAUM  
Children's stories—L. B. AMES

VOLUME 74—July-December, 1966

- Changes in personality for groups completing different amounts of college over two years—W. T. PLANT AND C. W. TELFORD  
Brain dynamism in perception of reversible figures—G. K. YACORZYNSKI  
Changes in Rorschach response throughout the human life span—L. B. AMES  
Mnemonic virtuosity: A study of chess players—A. BINET (Translated by M. L. SIMMEL AND S. B. BARRON)
- The development of behavior patterns in children—B. B. TYLER, F. B. TYLER AND J. E. RAFFERTY  
Correlates of classroom functioning—L. M. SMITH AND B. B. HIGGINS  
Children's behavior syndromes and parental responses—R. L. JENKINS, E. NUREDDIN, AND I. SHAPIRO

VOLUME 75—January-June, 1967

- Studies of an elementary game of strategy—B. SUTTON-SMITH AND J. M. ROBERTS  
Child care practices antedating three patterns of preschool behavior—D. BAUMRIND  
The impact of cognitive maturity on the development of sex-role attitudes in the years 4 to 8—L. KOHLBERG AND E. ZIGLER
- Differential factor structures in mathematical ability—P. S. VEAY  
Perception of the verbal as a function of affective stimuli and extraneous stimulation in schizophrenic and normal subjects—J. O. WELLS AND W. E. CALDWELL  
Mid-life attainment of the mentally retarded: A longitudinal study—W. R. BALLER, D. C. CHARLES, AND E. L. MILLER

VOLUME 76—July-December, 1967

- Scalogram analysis of some developmental sequences of infant behavior as measured by the Bayley Infant Scales of Mental Development—R. KOHEN-RAZ  
The leader and the lost: A case study of indigenous leadership in a poverty program community action committee—L. A. ZUCHER  
A developmental study of the perception of people—R. N. WOLMAN  
Human fears: A factor analytic study of three age levels—G. W. RUSSELL
- Correlates of voluntary social participation—J. E. TRELL  
Personality characteristics and developmental history of creative college women—R. HELSON  
A systematic study of therapeutic interventions in nursing homes—B. A. STOKSAY

VOLUME 77—January-June, 1968

- The fulfillment of promise: 40-year follow-up of the Terman Gifted Group—M. H. ODEN  
A scale for parents' ratings: Some ipsative and normative correlations—R. N. WALKER  
A methodological investigation of Piaget's concept of conservation of substance—A. C. BITTNER AND M. M. SHINEWELDO
- Sex differences in mental and behavioral traits—J. E. GARAI AND A. SCHEINFELD



# Genetic Psychology Monographs (continued)

## VOLUME 78—July-December, 1968

1. A theoretical and developmental investigation of delayed speech feedback—J. E. WATERS  
Growth changes in factorial structure by age and sex—N. W. DYE AND P. S. VEY  
A multispect investigation of masculine development in kindergarten age boys—H. B. BRILAN
2. The optimum limits of testing on the Wechsler Intelligence Scales—M. Y. QUEKESHI  
A developmental study of the ego-ideal—L. VAN DEN DAELE  
The relationship between physical condition immediately after birth and mental and motor performance at age four—N. EDWARDS

## VOLUME 79—January-June, 1969

1. Adolescents: From the views of the members of an informal adolescent group—T. P. JOSEPH  
Assaultation and childhood accidents among Chinese and Japanese Americans—M. KUROKAWA
2. Dimensions of achievement-related behavior among lower-class Negro parents—D. SOLOMON, R. J. PARELIUS, AND T. V. BUSSE  
Dimensions in the judgment of illness—J. S. KERNICK  
Reversal shifts and problem solving abilities at ages 4-7—S. B. KASTNER  
The judgment of body landmark heights—H. NASH  
Recognition of body surface regions—H. NASH

## VOLUME 80—July-December, 1969

1. Animal aggression: Review of research—W. M. VERNON  
The effects of positive and negative reinforcement and test anxiety on the reading performance of male elementary school children—S. CORLIE  
Personality and psychosomatic illness: A selective review of research—P. R. ROBBINS  
Task and motivational influences on the assessment of creative and intellectual ability in children—N. KOGAN AND F. T. MORGAN  
A study of an interview method for identifying family authority patterns—R. W. CHAMBERLIN
2. Conditioned reinforcing and aversive aspects of the stimuli defining the components of a two-component chain—A. KAUFMAN AND A. BARON  
The Henry Hilleman Center Family Interaction Scales—M. L. BENHENS, D. I. MYERS, W. GOLDFARB, N. GOLDFARB, AND N. D. FIELDSTEIN

## VOLUME 81—January-June, 1970

1. Attitudes of college students and their changes: A 37-year study—V. JONES  
Tattooing among male and female offenders of different ages in different types of institutions—A. J. W. TAYLOR
2. Sex differences in mental health—J. E. GARAI  
A selective review of dominance-subordination relations in animals—D. VAN KNEVELD  
The effects of visual or auditory stimulation given in infancy upon subsequent preference for stimulation of both modalities—W. R. BENKOWITZ  
The influence of certain child-rearing factors upon the development of a structured and salient sense of time—D. ZEIN

## VOLUME 82—July-December, 1970

1. Child care attitudes and emotional disturbance among mothers of young children—B. J. COHLER, J. L. WEISS, AND H. U. GRUNZBAUM  
A study of intragroup and intergroup relations among experimentally produced small groups—L. N. DIAN  
The development of role knowledge—D. R. HEISE AND E. P. M. ROBERTS
2. Middlelence: Age related stress periods during adult years—J. E. HORROCKS AND M. C. MUSSMAN  
The first six years of life: A report on current research and educational practice—E. LACROIX, P. C. LEE (Ed.), F. LITMAN, D. M. OSLIVIE, S. S. STODOLSKY, AND B. L. WHITE

## VOLUME 83—January-June, 1971

1. Measurement of parents' perceptions of their children's development—W. WOLFENBERGER AND R. A. KURTZ  
The concomitant development of cognitive and moral modes of thought: A test of selected deductions from Piaget's theory—L. C. LEE  
Children's influences upon parental behavior: An attempt to define the relationship with the use of laboratory tasks—J. D. OSOFSKY
2. Parent behavior and child academic achievement, achievement striving, and related personality characteristics—D. SOLOMON, K. A. HOULIHAN, T. V. BUSSE, AND R. J. PARELIUS  
Memory factors and school success of average and special groups of ninth-grade boys—M. MEKKER AND C. E. MEYERS  
Personality correlates of vocational planning—R. BARNETT

## VOLUME 84—July-December, 1971

1. Perceptual distortion as a function of stimulus objects, sex, naiveté, and trials using a portable model of the Ames Distorted Room—T. A. GERACE AND W. E. CALDWELL  
Calvanotic behavior and reinforcement of fish *Brachydanio Rerio*—W. F. FLANNAN, JR., AND W. E. CALDWELL  
Self-actualization, types of personality, and neuroses and psychoses—R. DREISTART  
Children's awareness of their bodily appearance: A developmental study of factors associated with the body percept—E. GILLERT, J. S. GARCUS, AND J. COHEN
2. Biomagnetic phenomena: Some implications for the behavioral and neurophysiological sciences—F. RUSSO AND W. E. CALDWELL  
Sociometric task leadership and the perception of deviance and subsequent conformity as functions of personality—W. L. LUBY, JR.  
Murderers who plead insanity: A descriptive factor analytic study of personality—M. W. KAHN

## VOLUME 85—January-June, 1972

1. Children's understanding of negation as a logical operation—S. S. FELDMAN  
Selective auditory perception of parental voices by the disturbed child—S. SANDERS  
Identity type, sociometric ratings, and group behavior—C. N. ZIMET AND C. SCHNEIDER
2. Patterns of perceived reaction to frustration among adolescents—J. O. PALMER  
The need for influence: Development and validation of a measure, and comparison with the need for power—J. S. ULKMAN  
C. Lloyd Morgan, E. C. Tolman, and comparative method—E. NEWSBURY  
Relearning motor skills in hemiplegia: Error analysis—L. DILLER, J. BUXBAUM, AND S. CHIOTELIS  
Effects of participation in rule formation on the moral judgment of children—R. L. MERCHANT AND F. REDELSKY  
Discrepancy and attention in the five-month infant—C. M. SUPER, J. KACAN, F. J. MORRISON, M. M. HAITE, AND J. WEIFFENBACH



# Genetic Psychology Monographs (*continued*)

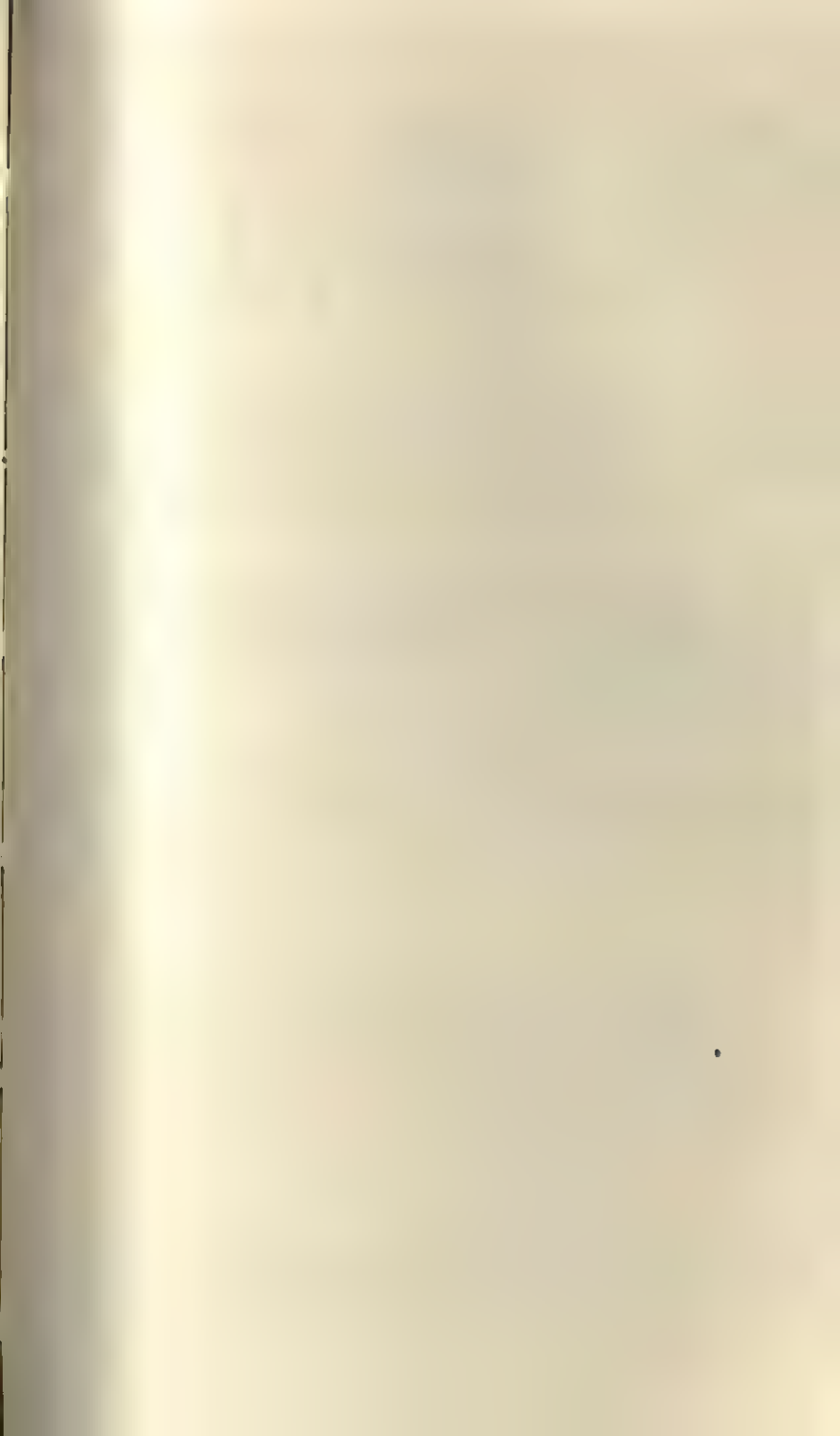
VOLUME 86—July-December, 1972

1. The role of enduring orientations to the environment in personality development—W. C. BRONSON  
Is there an error of the standard? A critique of Piaget's perceptual theory—D. A. BECKELMAN AND G. J. STEINFELD  
A cross-cultural study of the relationship between resistance to temptation and maternal communication—E. R. GREENGLASS  
The intellectual and achievement effects of preschool cognitive stimulation of poverty Mexican-American children—W. T. PLANT AND M. L. SOUTHERN
2. Factors affecting agnostic behavior in several subhuman species—R. F. GOTTIER  
Differential behavioral characteristics of nonpromoted children—J. A. CHASE  
An analysis of cognitive-developmental sequences—J. H. FLAVELL

VOLUME 87—January-June, 1973

1. A survey and interpretation of article characteristics: *The Journal of Genetic Psychology* 1945-1969—J. E. HONROCKS AND J. D. HOGAN  
The development of listener adapted communications in grade-school children from different social-class backgrounds—K. T. ALVY  
Locus of control among schizophrenics and other hospitalized psychiatric patients—T. F. CASH AND J. J. STACK  
Interactive styles and social adaptation—C. N. EDWARDS
2. The motivation of exploration strategies in young children—J. D. COSE  
Children's representation of the human figure: The effects of models, media, and instructions—C. GOLOMB  
A perceptual program for classroom teachers: Some results—R. N. WALKER AND J. STREFF  
Behavioral approaches to study and treatment of psychogenic tics—M. HERSEN AND R. M. EISLER  
Patterns of intellectual development during childhood and adolescence—M. Y. QUERREHI

THE JOURNAL PRESS  
2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.



## PUBLICATION IN THE JOURNAL OF PSYCHOLOGY

*The Journal of Psychology* provides a special service to authors and readers. Articles accepted for publication appear in print with the least possible delay. Although editorial standards are the same as those for the other journals, all correspondence and handling are expedited, and the referees may, at their discretion, permit some latitude in length of treatment with respect to tabular material, discussion, and review of the literature.

To save time, authors are requested to exercise unusual care in preparing their papers. Special attention should be paid to the General Instructions on the back cover of each issue and to the announcement on the first preliminary page.

\$39.00 per annum  
\$18.00 per volume  
Single Numbers \$9.00

BIMONTHLY  
Three volumes per year  
Immediate publication

July, 1973  
Volume 84, Second Half

(Founded in 1935 by Carl Murchison)

# The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a bilateral contract, and may not be resold for a period of five years from the date of publication.

JULY, 1973

(Manuscripts are printed in the order of final acceptance)

Sex-role identification and humor among preschool children . . . .	189
BY PAUL E. MCGHEE AND PHYLLIS GRODZITSKY	
The "lost letter technique" as a predictor of the 1972 presidential election .	195
BY MICHAEL JAY WEINER AND EDWARD LUREY	
A factor analysis of contraceptive preferences . . . . .	199
BY HARRISON G. GOUGH	
Parental dominance and sex-role identification in schizophrenia . . . .	211
BY NICHOLAS D. KOKONIS	
Psychological test performance and indigenous conceptions of intelligence .	219
BY ROBERT E. KLEIN, HOWARD E. FREEMAN, AND RICARDO MILLETT	
Signature size and self-estimation: A brief note . . . . .	223
BY E. R. MAHONEY	

(OVER)

Copyright © 1973 by The Journal Press  
Provincetown, Massachusetts, U. S. A.

All rights reserved

Second-class postage paid at Provincetown, Mass  
and additional mailing offices

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

Status of frustrator as a facilitator of aggression: A brief note . . . .	225
BY LAWRENCE J. CHASE AND NORBERT H. MILLS	
A political dimension on the I-E scale: A brief note . . . . .	227
BY MIRON ZUCKERMAN	
The development of personal space schemata toward body build . . . .	229
BY RICHARD M. LERNER	
Perception of environmental modifiability and involvement in antipollution activities . . . . .	237
BY HANNA LEVENSON	
Developmental trends in wishes, confidence, and the sense of personal control from childhood to middle maturity . . . . .	241
BY JOHN E. HORROCKS AND MILTON C. MUSSMAN	
The effect of first names on conflicted decisions: An experimental study . . .	253
BY THOMAS V. BUSSE AND CRAIG LOVE	
Language dominance and bilingual recall . . . . .	257
BY STARRETTE DALTON	
Scaling meaningfulness (M) of trigrams with children and retardates . . .	267
BY CECIL O. CAMPBELL AND ALFRED A. BAUMEISTER	
Practice effect in anagram solving . . . . .	279
BY EDWARD I. GAVURIN	
The effects of verbal reports of violence on aggression . . . . .	283
BY WILLIAM H. SCHARFF AND ROBERT S. SCHLOTTMANN	
Studies in the neurophysiology of learning: VIII. Oscillatory potentials resulting from cerebral self-stimulation in rats . . . . .	291
BY J. A. GENGERELLI	
Locus of control and self-disclosure of public and private information by college men and women: A brief note . . . . .	317
BY RICHARD M. RYCKMAN, MARTIN F. SHERMAN, AND GARY D. BURGESS	
MMPI two-point codes for a "normal" college male population: A replication study . . . . .	319
BY IVAN GULAS	
Some effects of fear of failure in the academic setting . . . . .	323
BY CARLOS GOLDBERG	
Intellectual ability as related to age and stage of disease in muscular dystrophy: A brief note . . . . .	333
BY F. WILLIAM BLACK	
The perceptions of parolees and parole officers . . . . .	335
BY EMILE S. SHIHADAH	
Concept-scale interaction with the semantic differential technique . . . .	345
BY DAVID L. KLEMMACK AND JOHN A. BALLWEG	
Effects of IQ and mental age on verbal imitative performance of children . .	353
BY REX FOREHAND, BRUCE ROBBINS, AND CHARLES PAT BRADY	



## SEX-ROLE IDENTIFICATION AND HUMOR AMONG PRESCHOOL CHILDREN\* <sup>1</sup>

*State University of New York at Albany*

---

PAUL E. MCGHEE AND PHYLLIS GRODZITSKY

---

### SUMMARY

The importance of mastery in children's humor has been emphasized by many researchers and theorists. The present study tested the hypothesis that a high level of mastery over S's sex-role identity enhances the funniness of humor based on sex inappropriate behavior. Seventeen three- to five-year-old boys were tested on the It Scale and were shown drawings of children engaged in sex appropriate or inappropriate behavior whose outcome was either positive or negative. While inappropriate and negative outcomes were consistently seen as funnier than appropriate and positive ones, when inappropriate positive outcomes were pitted against appropriate negative ones, boys choosing the former as funnier than the latter had significantly higher sex-role identity scores. Support was claimed for the general role of mastery in children's humor.

---

### A. INTRODUCTION

McGhee (6) reviewed the literature on children's humor and found both research and theory to point to the importance of mastery in young children's laughter and humor. That is, once a child has acquired a high level of conceptual mastery over some content area, he may perceive any inaccurate depiction of it as being funny. While studies dealing with this dimension have generally focused upon the role played by particular cognitive acquisitions [e.g., operational thinking (7, 8)] in the child's humor, no effort has been made to extend the investigation of mastery to broader personality dimensions. The present study conceptualizes sex-role identification in mastery terms and examines its role in determining the perception of humor in sex inappropriate behavior.

---

\* Received in the Editorial Office on March 26, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The authors would like to express their appreciation to Mrs. Adrian Shute, Director of the Robin Child Care Center, for making the subject population available for this study.

A number of studies have demonstrated that young children acquire sex-related behavioral preferences by the end of the preschool period (e.g., 1, 3, 5). Other findings (4, 9) suggest that sex-role identification patterns are already well established by age three. If the role played by violation of the child's concept of sex appropriate behavior is similar to that played by the violation of other more purely cognitive concepts, children with higher levels of mastery of sex appropriateness of various behaviors should find such violation funnier than children with lower mastery levels. The present study used preschool children to test this hypothesis.

## B. METHOD

### 1. Subjects

The Ss were 17 three- to five-year-old boys enrolled in an urban child care center in Albany, New York. All of these children came from lower class families, most of which were on welfare at the time the study was completed. Approximately half of the Ss tested were white, and half were black.

### 2. Stimulus Materials

The materials used consisted of the "It Scale" (1) and three sets of drawings. The drawings depicted children engaging in behaviors that were sex inappropriate for boys and sex appropriate for girls. These behaviors included (a) baking a cake, (b) putting on makeup, and (c) dressing a doll. The outcome (positive or negative) was also manipulated in these drawings (e.g., the cake turning out fine or exploding). Thus, four versions of each drawing were included, as follows:

BIP—boys engaged in sex inappropriate behavior with positive outcome.

BIN—boys engaged in sex inappropriate behavior with negative outcome.

GAP—girls engaged in sex appropriate behavior with positive outcome.

GAN—girls engaged in sex appropriate behavior with negative outcome.

### 3. Procedure

All Ss were tested individually, and the two different sets of stimulus materials were administered on two different days to each S. On the first day E (a female undergraduate student) administered the It Scale, using Brown's (1) procedure. On the second day S was shown pairs of drawings and was asked which one he thought was funnier and why. The E first randomly selected one of the three sets to show S. Within that set, two of the four drawings were randomly chosen for comparison; then the remaining two drawings of that set were presented to S. After randomly selecting a second

set of drawings, *E* repeated the above procedure for choosing drawings to be compared among the four possible remaining comparisons. This was again repeated for the two possible remaining comparisons for the third set of drawings. Thus, each *S* made two choices of the funnier drawing for each of the three content areas (making a cake, putting on makeup, and dressing a doll). No drawing was presented twice to any *S*. Each *S*, therefore, made six comparisons in random order: 1) BIP *vs.* BIN, 2) GAP *vs.* GAN, 3) BIN *vs.* GAP, 4) BIP *vs.* GAN, 5) BIP *vs.* GAP, and 6) BIN *vs.* GAN. While two comparisons kept sex appropriateness of the behavior constant and varied outcome, and two kept outcome constant and varied sex appropriateness, the remaining two varied both outcome and sex appropriateness.

Before presenting any of the drawings, *E* said: "Now I'm going to tell you some stories about children who like to play in their house and show you some pictures of them." In each case, after presenting the two drawings *E* asked *S* which one was funnier and why.

### C. RESULTS

To determine the relationship between choice of the funnier drawing and sex-role identity, six *t* tests were computed on *Ss'* It Scale scores as a function of drawing chosen. Only the test for BIP *vs.* GAN was significant ( $p < .01$ ), with *Ss* choosing BIP as funnier having significantly higher sex-role identification than *Ss* choosing GAN. A binomial test (10) was computed to determine the probability of obtaining the frequency of choices obtained in specific funniness comparisons. The following results were obtained: 1) BIP *vs.* BIN, 76% BIN ( $p < .025$ ); 2) GAP *vs.* GAN, 70% GAN ( $p < .072$ ); 3) BIN *vs.* GAP, 82% BIN ( $p < .006$ ); 4) BIP *vs.* GAN, 59% GAN ( $p < .166$ ); 5) BIP *vs.* GAP, 82% BIP ( $p < .006$ ); 6) BIN *vs.* GAN, 50% BIN ( $p < .598$ ).

### D. DISCUSSION

Before consideration of the significance of the BIP *vs.* GAN comparison, the general trends found among the remaining comparisons should be noted. There are two possible sources of humor in the drawings used: the sex inappropriateness of the behavior and the negativity of the outcome, each of which requires some level of cognitive mastery over the stimulus content in order to become a potential basis for humor. A high level of familiarity with or mastery over the sex-linked behaviors depicted generates the expectation in *Ss* that only girls will engage in these behaviors and the outcome is predictable (the "right" outcome being having makeup on straight, etc.). Thus,

violation of either of these expectations provides an opportunity for Ss to perceive the content as humorous. This may be seen above in the predominance of choices (as funnier) of drawings depicting inappropriateness or a negative outcome over those depicting appropriateness of a positive outcome. When a drawing contains both appropriate behavior and a positive outcome (GAP), it clearly is not often chosen as the funnier drawing (comparisons 2, 3, and 5). The fact that comparison 3 reached the highest level of significance is not surprising, since it is the only comparison putting a drawing with two possible sources of humor against a drawing with none. The fact that the same percentage of Ss chose BIP in comparison 5 as chose BIN in comparison 3 suggests that the inappropriateness of the behavior may play a more important role than its negative outcome in determining Ss' choices of the funnier drawing. The findings for comparison 6, however, are inconsistent with this interpretation. When negativity of outcome is held constant, choices of appropriate *versus* inappropriate behaviors as funnier are evenly split. Thus, it is difficult to separate the differential contributions of appropriateness and outcome to funniness of the drawings.

Only comparison 4 (BIP *vs.* GAN) pitted the two bases for humor against each other, with no clear preference being shown for either drawing. However, it was only when a choice had to be made between these two that the influence of level of sex-role mastery became evident. Boys with high sex-role identification were more likely to choose BIP as funnier, while boys with low sex-role identification tended to choose GAN as funnier. Wolfenstein (11) has suggested that children begin to laugh at "name change" statements ("Hi, Sally," to a boy) or "gender change" statements ("You're a girl," to a boy) only when they become confident of their own sexual identity. Prior to achieving such confidence, their reaction is likely to be anxiety and concern rather than laughter. In light of Coopersmith's (2) and others' finding that low self-concept children are more likely to be anxious and defensive, it may be that Ss in the present study with lower sex-role identification saw the sex inappropriate behavior in comparison 4 as less funny because of anxiety over their own sexual identity. While the role of anxiety in these data cannot be clearly determined in this study, the findings do support the hypothesis that higher levels of sex-role mastery play an important role in determining appreciation of humor based on sex inappropriate behavior.

#### REFERENCES

1. BROWN, D. G. Sex-role preference in young children. *Psychol. Monog.*, 1956, 70(14), Whole No. 421.
2. COOPERSMITH, S. *The Antecedents of Self-Esteem*. San Francisco: Freeman, 1967.

3. DELUCIA, L. A. The toy preference test: A measure of sex-role identification. *Child Devel.*, 1963, **34**, 107-117.
4. HARTUP, W. W., & ZOOK, E. A. Sex-role preferences in three- and four-year-old children. *J. Consult. Psychol.*, 1960, **24**, 420-426.
5. KAGAN, J. Acquisition and significance of sex typing and sex-role identity. In M. L. Hoffman & L. W. Hoffman (Eds.), *Review of Child Development Research (Vol. 1)*. New York: Russell Sage Foundation, 1964. Pp. 137-167.
6. MCGHEE, P. E. Development of the humor response: A review of the literature. *Psychol. Bull.*, 1971, **76**, 328-348.
7. ———. Cognitive development and children's comprehension of humor. *Child Devel.*, 1971, **42**, 123-138.
8. ———. On the cognitive origins of incongruity humor: Fantasy assimilation versus reality assimilation. In J. H. Goldstein & P. E. McGhee (Eds.), *The Psychology of Humor*. New York: Academic Press, 1972. Pp. 61-80.
9. SCHELL, R. E., & SILBER, J. W. Sex-role discrimination among young children. *Percept. & Motor Skills*, 1968, **27**, 379-389.
10. SIEGEL, S. *Nonparametric Statistics*. New York: McGraw-Hill, 1956.
11. WOLFENSTEIN, M. *Children's Humor*. Glencoe, Ill.: Free Press, 1954.

*Psychology Department*

*State University of New York at Albany*

*1400 Washington Avenue*

*Albany, New York 12222*





## THE "LOST LETTER TECHNIQUE" AS A PREDICTOR OF THE 1972 PRESIDENTIAL ELECTION\*

*University of North Carolina at Greensboro*

---

MICHAEL JAY WEINER AND EDWARD LUREY

---

### SUMMARY

The "lost letter technique" was validated against actual election returns in two randomly selected precincts in Greensboro, North Carolina. The technique was not successful in predicting the outcome of the presidential race. It was found that letters addressed to The Health Research Center were significantly more likely to be returned than letters addressed to presidential candidates. This indicated that the failure to predict was due either to voter apathy in this particular election or to an active dislike of both candidates.

---

### A. INTRODUCTION

Milgram (1) devised the "lost letter technique" to alleviate the problem of reactivity in survey research. The technique required that posted letters be left in a designated area where they could be found by people passing by. The letters were addressed to a committee advocating or opposing a specific position or candidate or to a neutral third party. The difference in the return rate of the lost letters was used as a measure of the prevailing attitude in the area.

The technique was found to be successful in predicting attitudes toward the Nazi and Communist parties, integration in southern states, and the results of the Johnson-Goldwater election in 1964 (1).

An attempt to validate the technique against actual election returns from two randomly selected precincts in Greensboro, North Carolina, was made prior to the recent presidential election.

### B. METHOD

The letters were distributed at night between October 23rd and November 3rd by eight undergraduates enrolled in the social psychology course at the

---

\* Received in the Editorial Office on April 4, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

University of North Carolina at Greensboro. They were distributed in two precincts which were selected at random from the 10 precincts with the highest rate of voter registration. Each envelope was sealed and stamped, and each contained an appropriately worded ambiguous letter. In each precinct, 125 letters were addressed to Concerned Citizens To Re-elect Nixon, 125 to Concerned Citizens To Elect McGovern, and 50 to The Health Research Center. All of the letters were addressed to the same post office box rented specifically for the study.

Each letter, along with a handwritten note which read "found near car," was placed under the windshield of a car selected by a random process. Milgram (1) found the technique most accurate when letters were distributed in this manner.

### C. RESULTS AND DISCUSSION

The results are presented in Table 1. It is evident that the technique was not successful in predicting the outcome of the election ( $\chi^2 = .13$ ). Although Milgram (1) found that the technique was likely to underestimate the extent of actual election returns, in the present study even the trend was not predicted.

There was a significantly greater return rate in Precinct 2 than in Precinct 1 ( $\chi^2 = 11.79$ ,  $p < .001$ ) which was probably due to a socioeconomic class difference. All of the dwellings in Precinct 2 were private and had among the highest value in Greensboro. Precinct 1 consisted of small private homes and apartments.

The results also indicated that letters addressed to The Health Research Center had a significantly higher return rate than letters supporting presidential candidates ( $\chi^2 = 21.22$ ,  $p < .001$ ). This effect was significant in Precinct 1 ( $\chi^2 = 8.36$ ,  $p < .005$ ) and in Precinct 2 ( $\chi^2 = 14.71$ ,  $p < .001$ ). Since Milgram found that there were no differences in return rates for letters addressed to an unknown party and those addressed to a medical research

TABLE 1  
PERCENTAGE OF LETTERS RETURNED AS COMPARED TO  
ACTUAL ELECTION RETURNS

Candidate	n	Precinct 1		n	Precinct 2	
		% returned	% of vote		% returned	% of vote
McGovern	125	37	19	125	48	24
Nixon	125	38	79	125	51	75
Control	50	58		50	78	

organization, the results might indicate an apathetic attitude toward the election.

On the other hand, as Milgram contended, failure to mail a letter could be conceived of as an active act rather than a passive one. If this was the case, the depressed return rate could have indicated a general dislike for both of the candidates, rather than apathy. The reason for the failure to predict could only have been answered by follow-up research which was beyond the resources of the present investigation.

#### REFERENCES

1. MILGRAM, S. The lost letter technique. *Psychol. Today*, 1969, **3**, 30.

*Department of Psychology*

*The University of North Carolina at Greensboro*

*Greensboro, North Carolina 27412*





## A FACTOR ANALYSIS OF CONTRACEPTIVE PREFERENCES\* 1

*University of California, Berkeley*

---

HARRISON G. GOUGH

---

### SUMMARY

Factor analysis of acceptability ratings given to 10 contraceptive methods by 178 Ss (78 male, 100 female) identified four groupings: (a) coitus-dependent—condom, diaphragm, and foam or jelly; (b) surgical—vasectomy and tubal ligation; (c) coitus-inhibiting—abstinence, rhythm, and withdrawal; and (d) coitus-independent—intrauterine device and pill. Factor scores were computed and correlated with variables from the Adjective Check List, California F Scale, California Psychological Inventory, College Vocabulary Test, Miller-Fisk Sexual Knowledge Questionnaire, and Rotter's locus-of-control scale. Although the yield of significant correlations was modest, several heuristic patterns of covariation with the personality and cognitive measures were noted.

---

### A. USE OF CONTRACEPTIVE METHODS

In the last 30 years there has been a steady increase in the number of American men and women who believe that birth control and contraceptive information should be generally available. For example, in 1937 polls indicated that 66% of males and 70% of females held this attitude, whereas by 1969 the percentages had gone up to 89 and 86 (5). Usage of contraceptive methods has similarly increased, with the proportion of fecund women in the United States who have used or plan to use contraception recently estimated at 97% (22). Similar trends are discernible on a worldwide basis (4), and in many countries the governments themselves have adopted explicit policies in regard to family planning and population control.

#### 1. Choice Among Methods

Choice of contraceptive method is one of the researchable topics that lies within this vast nexus of social change and development. Of the 12 or more

---

\* Received in the Editorial Office on April 9, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

1 Work on this study was supported by grant No. NIH-HD-06589-01 from the National Institute of Child Health and Human Development, United States Public Health Service.

known methods of contraception, none is foolproof and all involve risks of one sort or another (16). Choice of technique will therefore be based on the medical considerations in a particular case, personal preference, or even myth and dieristic thinking (3, 18).

The most recent findings from the 1970 National Fertility Survey (21) show the pill to be first in popularity among contraceptive methods, with approximately 34% of American couples employing this technique. Surgical sterilization of husband was reported by 7.8% of the couples, and of wife by 8.5%. The condom was next in frequency, being reported by 14.2% of the couples surveyed. In diminishing order of usage came the intrauterine (IUD) device (7.4%), rhythm method (6.4%), foam (6.1%), diaphragm (5.7%), douche (3.2%), and withdrawal (2.1%). It might be noted that about 65% of the 3810 couples surveyed were using one or more forms of contraception, whereas 35% were employing none.

Surgical methods of contraception are being used with increasing frequency. In 1964 Campbell (6) estimated that approximately 45,000 vasectomies were being performed per year, but by 1972 Davis (7) estimated the yearly total at 700,000. Westoff (21) suggested that as of 1970 some 2.75 million couples of reproductive age had resorted either to tubal ligation or vasectomy. Among older couples (wife aged 30 to 44) practicing contraception, surgical methods are now the most frequently used techniques, with percentages of 13.1 for ligation and 12.1 for vasectomy. This increase is of interest when set against medical opinion (2) that tubal ligation should be reserved for cases where prevention of pregnancy is essential for either emotional or physical reasons.

## *2. Changes from One Method to Another*

An interesting observation made in recent years is that Ss tend to change contraceptive methods, and there may even be a sort of trial-and-error learning process that goes on before a couple settles on the preferred technique. Moos (12) noted that 40% of his sample of 718 women had changed methods prior to being included in his study, and 20% indicated that they were contemplating a change in the future. Murawski (13) found that 35% of his Ss shifted to a different form of oral contraception from the one used at the start of the study, and that 36% later gave up the pill in favor of the IUD. The "pill scare" of 1969-70 (14) resulted in perhaps 1.5 million women at least temporarily abandoning the pill, leading to what Ziegler (23) called a "major epidemic of unwanted pregnancies."

Given the variety of contraceptive methods available and the lack of com-

elling evidence on the superiority of any one, wide latitude is allowed to individual preferences. The question can then be asked, "Are there any underlying clusters or dimensionalities among these choices, such that persons favoring one method will have predictable patterns of acceptances and rejections among those remaining?" It is the purpose of this paper to search for such patterns of preference among 10 widely used contraceptive methods, employing the technique of factor analysis.

## B. RESEARCH SAMPLES AND PROCEDURES

A list of 10 contraceptive methods was drawn up (see Table 1)<sup>2</sup> and presented in a five-step rating format, going from 5 = very acceptable, 4 = somewhat acceptable, 3 = not sure, 2 = somewhat unacceptable, to 1 = very unacceptable. A sample of 78 male Ss was composed of 25 men who were participating in a community study of environmental attitudes, nine other adults tested on an individual basis by the author, and 44 college students who completed the inquiry as part of their outside laboratory work in an intermediate course in psychology. Mean age for the adults was 41.91,  $SD = 12.91$ , and for the students 20.05,  $SD = 2.88$ ; for all 78 males, mean age was 29.58,  $SD = 13.88$ .

The 100 females tested included 25 from the community survey, 35 mothers with children in a cooperative nursery school, and 40 college students. Mean age for the adults was 34.82,  $SD = 10.28$ , and for the students 19.95,  $SD = 2.02$ ; for all 100 Ss, mean age was 28.87,  $SD = 10.82$ .

The questionnaires were filled in anonymously by the students, to avoid any embarrassment and to insure confidentiality of reply, although test forms were coded so that one measure could be collated with another. The adults were tested under a more personal arrangement, and names were

---

<sup>2</sup> Two contraceptive methods that may be used with increased frequency in future practice were excluded from the list, as in pretesting it was found that very few subjects were familiar with them. Injection of ovulation suppressing substances (as opposed to the more common oral ingestion of the "pill") was one of these. The other was what is probably the most recent addition to the armamentarium of contraceptive agents, the acidic lipid called "prostaglandin" by von Euler (19, 20). In the 1960's it was noted by Karim (11) that prostaglandins were present in high concentration in women at or near labor. Later he found that infusion of biosynthesized prostaglandins would induce labor in parturient women and that the drugs would also induce abortion in gravid Ss. Vaginal suppositories administered a day or two before normal menses are currently being evaluated. Although the prostaglandins may prove to be the ideal "morning after" or "month after" contraceptive, they have not been under study long enough at this time to determine contraindications, side effects, and the like. Speroff and Ramwell (17) have published a detailed analysis of currently available information on the approximately 13 human prostaglandins so far identified.

TABLE 1  
CORRELATIONAL AND NORMATIVE ANALYSES OF 10 CONTRACEPTIVE METHODS

Method	1	2	3	4	5	6	7	8	9	10	Normative data	
											M	SD
1. Condom	—	.52	.34	.06	.07	.12	.18	.18	.06	.04	3.35	1.21
2. Diaphragm		—	.45	.08	.06	.08	.10	.17	.34	.20	3.83	1.10
3. Foam or jelly			—	-.02	.04	-.03	.06	.02	.26	.04	3.39	1.14
4. Vasectomy				—	.82	-.07	-.33	-.07	.13	.20	3.76	1.39
5. Tubal ligation					—	-.01	-.33	-.04	.22	.23	3.58	1.37
6. Abstinence						—	.38	.36	-.15	-.04	1.48	.94
7. Rhythm							—	.36	-.24	-.24	2.41	1.28
8. Withdrawal								—	.04	-.01	1.81	1.04
9. IUD									—	.38	4.06	1.11
10. Pill										—	4.25	1.08

Note: Based on a sample of 178 Ss (78 males, 100 females).

therefore recorded; individual confidentiality of protocols was, of course, assured to the adults as well as to the students.

### C. FINDINGS

Table 1 presents the intercorrelations among the ratings given to the 10 contraceptive methods by the total sample of 178 Ss, and also the mean ratings for each method. Attention is centered on findings for the total sample of 178 Ss because prior analyses of the ratings for each sex separately had revealed only minor differences. For example, the rank-difference correlation between rankings of males and females was  $+.94$ , and the only significant contrast by method was for the pill, where males gave a slightly more favorable rating than did females (difference =  $.33$ ,  $t = 2.02$ ). Both males and females assigned their two highest preferences to the pill and intrauterine device or loop, and both gave their lowest acceptability ratings to abstinence. For the total sample, the rank-order of preference was as follows: pill (greatest acceptability), IUD, diaphragm, vasectomy, tubal ligation, condom, foam or jelly, rhythm method, withdrawal, and abstinence.<sup>8</sup> Factor analyses of the separate  $10 \times 10$  matrices for males and females also gave nearly equivalent results.

Table 2 gives the findings from a principal axes factor analysis before and after varimax rotation with communalities set at unity. The first factor after rotation had major loadings on the condom, diaphragm, and foam or jelly, methods that are all coitus-dependent. Factor 2 had major loadings on vasectomy and tubal ligation, and is clearly a surgical-sterilization factor. Factor 3 includes abstinence, the rhythm method, and withdrawal. Abstinence and the rhythm method both involve a refraining from intercourse, and withdrawal a cessation. This common theme of inhibition of behavior may be posited as the basic element in Factor 3. Factor 4, with major loadings on the IUD and pill, may be designated as a coitus-independent factor.

Considerations such as these permit suggestion of the following names for each factor: Factor 1, coitus-dependent; Factor 2, surgical; Factor 3, coitus-inhibiting; and Factor 4, coitus-independent. The covariations expressed in the factor loadings imply that an individual changing from one mode of contraception to another would be more likely to change to a different technique within the same factor than to a technique classified under another factorial heading. For example, a woman deciding to change from the pill would be

<sup>8</sup> The rank-order correlation between this sequence and rankings based on frequency of usage as reported by Westoff (21) for the National Fertility Survey of 1970 was  $+.65$ .



TABLE 2  
FACTOR ANALYSIS OF 10 CONTRACEPTIVE METHODS

Method	Factor loadings							
	Before rotation				After varimax rotation			
	1	2	3	4	1	2	3	4
1. Condom	.71	.16	.00	-.37	.77	.12	.20	-.10
2. Diaphragm	.76	.33	-.20	-.07	.79	.02	.14	.28
3. Foam or jelly	.58	.24	-.38	-.29	.77	-.07	-.13	.08
4. Vasectomy	-.14	.74	.52	-.24	.03	.94	-.07	.07
5. Tubal ligation	-.09	.76	.52	-.17	.04	.93	-.03	.15
6. Abstinence	.40	-.30	.59	.20	-.02	.05	.79	-.10
7. Rhythm	.49	-.60	.22	-.01	.18	-.35	.62	-.32
8. Withdrawal	.52	-.21	.44	.36	.08	-.04	.78	.14
9. IUD	.25	.58	-.32	.45	.25	.05	-.12	.79
10. Pill	.09	.54	-.07	.61	-.02	.16	.03	.80
% variance	.22	.24	.14	.11	.19	.19	.17	.15

expected to think first of the IUD; one using foam or jelly would think first of changing to a diaphragm, or condom for her male partner; a man considering vasectomy would also be favorably disposed toward tubal ligation of his wife as an alternate method of contraception.

Factor scores were next computed for each of the 178 Ss, assigning each contraceptive method to the single factor on which its rotated loading was highest. Computations, that is to say, were based on the following equations: coitus-dependent = .77 condom + .79 diaphragm + .77 foam; surgical = .94 vasectomy + .93 ligation; coitus-inhibiting = .79 abstinence + .62 rhythm + .78 withdrawal; coitus-independent = .79 IUD + .80 pill.

Means and sigmas for each sex on the four factor scores are given in Table 3. Although variations occur, no difference between means is significant at the .05 level of confidence. Intercorrelations among the four factor scores were as follows: 1 vs. 2,  $r = .09$ ; 1 vs. 3,  $r = .28$ ; 1 vs. 4,  $r = .36$ ; 2 vs. 3,  $r = -.33$ ; 2 vs. 4,  $r = .54$ ; and 3 vs. 4,  $r = -.33$ .

The adult and student subsamples mentioned earlier were also compared on the four factor scores. For the males, means for the 34 adults and 44 students

TABLE 3  
COMPARISON OF 78 MALES AND 100 FEMALES ON FOUR CONTRACEPTIVE FACTORS

Factor	Males		Females		diff	t
	M	SD	M	SD		
Factor 1	10.20	2.23	9.76	2.27	.44	1.28
Factor 2	6.77	2.86	7.43	2.57	-.66	1.61
Factor 3	4.39	2.10	3.90	1.75	.49	1.69
Factor 4	7.79	1.79	7.72	1.90	.07	.24

were as follows: Factor 1, 10.29 *vs.* 10.12; Factor 2, 6.47 *vs.* 7.01; Factor 3, 4.10 *vs.* 4.61; and Factor 4, 7.75 *vs.* 7.83. None of these differences is significant at  $p = .05$ . For the females, means for the 60 adults and 40 students were as follows: Factor 1, 9.50 *vs.* 10.15; Factor 2, 7.97 *vs.* 6.62; Factor 3, 3.25 *vs.* 4.86; and Factor 4, 7.56 *vs.* 7.97. The differences here on Factors 2 and 3 are significant ( $p < .01$ ). The adult women gave higher acceptability ratings to the surgical procedures, and the students gave higher ratings to the coitus-inhibiting methods.

#### D. INTERPRETATIONAL DATA

The 178 Ss in the study were also given the Adjective Check List—ACL (9), California F Scale for authoritarianism (1), California Psychological Inventory—CPI (8), Rotter's scale for locus of control—LOC (15), College Vocabulary Test—CVT (10), and an unpublished 24-item test of sexual knowledge developed by Miller and Fisk at the Stanford University School of Medicine.<sup>4</sup> Age was also analyzed, in view of Westoff's report (21) showing a positive correlation between age and use of surgical methods, and a negative correlation between age and use of the pill.

The correlations obtained between the four contraceptive preference factors and the other variables are given in Table 4. Because of the rather modest yield of significant relationships in the complete table of 368 coefficients, only those that appear to warrant discussion will be mentioned in the text. For the male sample, correlations  $\geq .22$  and  $\geq .29$  are significant at the .05 and .01 levels of probability, respectively, and the corresponding coefficients for the female sample are  $\geq .20$  and  $\geq .26$ .

Review of the correlations computed for the coitus-dependent Factor 1 identified only minimal trends: correlations of  $-.13$  and  $-.21$  for males and females with the self-control scale of the CPI,  $-.06$  and  $-.20$  with the CPI scale for sense of well-being, and  $-.09$  and  $-.08$  with the autonomy scale of the ACL. Personality and intellectual variables, as assessed on these Ss, do not appear to be appreciably related to favorability of response to the three coitus-dependent contraceptive methods.

A somewhat more encouraging yield was found for Factor 2. The authoritarian personality scale correlated  $-.34$  for males and  $-.18$  for females, and the sexual knowledge test gave coefficients of .25 and .11. Age had the expected positive correlation for females ( $+.18$ ), but produced a negative

<sup>4</sup> The writer is indebted to Drs. Warren Miller and Norman Fisk of the Department of Psychiatry, Stanford University School of Medicine, for their kindness in making this test available.

TABLE 4  
CORRELATIONS BETWEEN FOUR CONTRACEPTIVE FACTOR SCORES AND THE VARIABLES  
INDICATED, IN SAMPLES OF 78 MALES AND 100 FEMALES

Variable	Factor 1		Factor 2		Factor 3		Factor 4	
	M	F	M	F	M	F	M	F
<b>A. Adjective Check List</b>								
1. Number checked	-.13	.12	-.11	-.02	.02	.11	-.05	.03
2. Defense	-.05	-.02	-.12	.10	.02	-.05	-.13	-.06
3. Favorable	-.06	.06	-.18	.05	-.02	.03	-.07	-.05
4. Unfavorable	-.01	.03	.09	.00	.01	.14	.10	.11
5. Self-confidence	-.01	.01	.04	.12	-.19	-.11	.20	.22
6. Self-control	-.03	-.03	-.26*	.08	.13	.01	-.26*	-.19
7. Liability	.05	.04	.20	-.07	-.16	-.04	.16	.03
8. Personal adjustment	-.01	-.06	-.12	.03	.00	-.04	-.18	-.14
9. Achievement	-.06	.04	-.02	.13	-.22*	-.02	.05	.04
10. Dominance	-.05	-.09	.05	.11	-.23*	-.08	.17	.12
11. Endurance	-.01	-.02	-.20	.06	-.02	.01	-.16	-.12
12. Order	.00	-.02	-.19	.11	-.05	-.02	-.13	-.10
13. Intraception	-.01	-.01	-.16	.11	-.02	-.06	-.13	-.11
14. Nurturance	.05	.00	-.19	.02	.06	-.08	-.16	-.10
15. Affiliation	-.11	.04	-.16	-.01	.11	.01	-.17	-.06
16. Heterosexuality	.13	-.05	-.01	-.10	.07	-.07	.12	.10
17. Exhibition	.01	-.10	.08	-.04	-.13	-.05	.17	.20*
18. Autonomy	-.09	-.08	.15	-.02	-.20	-.03	.22*	.14
19. Aggression	-.04	-.11	.18	.05	-.18	-.09	.20	.18
20. Change	-.06	.12	.16	-.02	-.23*	.06	.20	.14
21. Succorance	.04	.11	-.05	-.16	.15	.11	-.18	-.07
22. Abasement	.02	.16	-.20	.02	.18	.04	-.26*	-.10
23. Deference	.10	.06	-.18	-.03	.20	.00	-.20	-.22*
24. Counseling readiness	.00	-.14	.05	-.01	.16	.03	-.11	-.02
<b>B. California Psychological Inventory</b>								
1. Dominance	.05	-.09	-.12	.10	-.02	-.08	.10	.00

TABLE 4 (continued)

Variable	Factor 1		Factor 2		Factor 3		Factor 4	
	M	F	M	F	M	F	M	F
2. Capacity for status	.01	-.05	.10	.15	-.33**	-.22*	.12	.03
3. Sociability	.10	-.04	-.03	.02	-.15	-.06	.13	-.03
4. Social presence	.04	-.04	-.09	.06	-.13	-.09	.03	.11
5. Self-acceptance	.09	.01	-.13	.03	-.04	-.07	.14	.03
6. Sense of well-being	-.06	-.20*	.14	.06	-.16	-.20*	.10	-.12
7. Responsibility	.10	-.06	.03	-.03	.01	-.11	.03	-.25*
8. Socialization	.12	-.08	.10	.01	.20	.07	.09	-.23*
9. Self-control	-.13	-.21*	.01	-.04	-.08	-.11	-.15	-.27**
10. Tolerance	-.03	-.10	.07	.08	-.11	.14	.00	-.18
11. Good impression	-.18	-.11	.05	-.02	-.12	-.01	-.14	-.18
12. Communality	.19	-.04	.00	-.03	-.12	-.08	.17	-.16
13. Achievement via conformance	.05	-.18	.08	-.07	.02	-.14	.00	-.29**
14. Achievement via independence	.05	-.06	.03	.03	-.06	-.26**	.01	-.08
15. Intellectual efficiency	.11	-.19	.00	.04	-.07	-.21*	.06	-.11
16. Psychological-mindedness	-.05	-.05	-.15	.21*	-.10	-.36**	-.16	.04
17. Flexibility	-.08	.00	.03	.09	-.02	-.14	-.08	.10
18. Femininity	.03	-.09	-.14	.17	.01	-.24*	-.12	-.01
C. California F Scale (authoritarianism)	-.39**	.14	-.34**	-.18	-.08	.41**	-.31**	-.05
D. College Vocabulary Test	.13	.16	.17	.08	.02	-.08	-.16	.02
E. Locus of control (external control)	.13	.06	.17	-.04	.02	.01	.17	.14
F. Sexual Knowledge Questionnaire	.22*	-.14	.25*	.11	-.24*	-.55**	.42**	.15
G. Age	.10	-.06	-.04	.18	-.11	-.13	.00	-.12

\*  $p < .05$ .\*\*  $p < .01$ .

value ( $-.04$ ) among males. No trends were observed on the ACL or CPI, or on the locus of control variable. Tentative inferences may be drawn that persons with favorable attitudes concerning surgical methods of contraception are less authoritarian in political and social outlook, and better informed about sexual matters than persons who give surgical methods a low acceptability rating.

Factor 3, coitus-inhibiting methods, correlated  $-.33$  for males and  $-.22$  for females with the capacity for status scale of the CPI,  $-.16$  and  $-.20$  with CPI sense of well-being,  $-.10$  and  $-.36$  with CPI psychological-mindedness,  $-.11$  and  $-.13$  with age,  $-.23$  and  $-.08$  with ACL dominance, and  $-.24$  and  $-.55$  with the sexual knowledge test. With due caution, it can be suggested that Ss giving a high acceptability rating to Factor 3 methods lack the sense of self-confidence and well-being found among those who rate Factor 3 methods low, are less optimistic concerning their future prospects, and less well informed concerning childbirth, conception, and related issues.

The critical variables for Factor 4 appear to be self-control on the CPI with correlations of  $-.15$  and  $-.27$  for males and females, good impression (coefficients of  $-.14$  and  $-.18$ ), sexual knowledge (.42 and .15); and on the ACL the scales for self-confidence (.20 and .22), self-control ( $-.26$  and  $-.19$ ), exhibition (.17 and .20), autonomy (.22 and .14), aggression (.20 and .18), change (.20 and .14), abasement ( $-.26$  and  $-.10$ ), and deference ( $-.20$  and  $-.22$ ). Those responding more favorably to the IUD and pill score higher on the measure of sexual knowledge, express impulse and aggression more easily, seek and enjoy change, and react to others in a less deferential manner. Although the magnitude of these relationships is low, their abundance suggests that a rather good forecast of higher or lower ratings of Factor 4 methods could be made on the basis of a multiple regression analysis of the auxiliary personality and cognitive variables.

It should be emphasized that the resumés attempted above are impressionistic and highly speculative. They should be viewed as tentative formulations, awaiting confirmation or disproof on new and larger samples. The only conclusion warranted at this time is that factorial clusterings do appear to exist among the various contraceptive methods now available. Shifts from one technique to another, therefore, should follow predictable routes; that is, they should not be mere matters of chance. The cognitive and personality data adduced in the last section of the paper hint that these preferences in the contraceptive domain may be aligned with more general patterns of psychological variation.



## REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWICK, E., LEVINSON, D. J., & SANFORD, R. N. *The Authoritarian Personality*. New York: Harper, 1950.
2. BAKKER, C. B., & DIGHTMAN, C. R. Physicians and family planning: A persistent ambivalence. *Obstet. & Gynec.*, 1965, **25**, 279-284.
3. BALFOUR, M. C., CHANDRASEKARAN, C., GUTTMACHER, A. F., KOYA, Y., KOYA, T., & RAINWATER, L. Panel discussion on comparative acceptability of different methods of contraception. In C. V. Kiser (Ed.), *Research in Family Planning*. Princeton: Princeton Univ. Press, 1962, Pp. 373-386.
4. BAUMGARTNER, L. Family planning around the world. In B. Berelson, R. K. Anderson, O. Harkavy, J. Maier, W. P. Mauldin, & S. J. Siegel (Eds.), *Family Planning and Population Programs: A Review of World Developments*. Chicago: Univ. Chicago Press, 1966. Pp. 277-294.
5. BLAKE, J. Population policy for Americans: Is the government being misled? *Science*, 1969, **165**, 367-373.
6. CAMPBELL, A. A. The incidence of operations that prevent conception. *Amer. J. Obstet. & Gynec.*, 1964, **89**, 694-700.
7. DAVIS, J. E. Vasectomy. *Amer. J. Nurs.*, 1972, **72**, 509-513.
8. GOUGH, H. G. *Manual for the California Psychological Inventory*. Palo Alto, Calif.: Consult. Psychol. Press, 1957.
9. GOUGH, H. G., & HEILBRUN, A. B., Jr. *The Adjective Check List Manual*. Palo Alto, Calif.: Consult. Psychol. Press, 1965.
10. GOUGH, H. G., & SAMPSON, H. *The College Vocabulary Test*. Berkeley: Univ. California Inst. Personal. Assess. & Res., 1954.
11. KARIM, S. M. M. Identification of prostaglandins in human amniotic fluid. *J. Obstet. & Gynec. Brit. Commonwealth*, 1966, **73**, 903-908.
12. MOOS, R. H. Psychological aspects of oral contraceptives. *Arch. Gen. Psychiat.*, 1968, **19**, 87-94.
13. MURAWSKI, B. J. Psychologic considerations for the evaluation of long-term use of oral contraceptives. *J. Reprod. Med.*, 1969, **3**, 151-155.
14. RORVIK, D. M. Beyond the pill. *Look*, 1971, **35**(12), 17-20.
15. ROTTER, J. B. Generalized expectancies for internal versus external control of reinforcement. *Psychol. Monog.*, 1966, **80**(1), Whole No. 609.
16. SOUTHAM, A. L. Contraceptive methods: Use, safety, and effectiveness. In B. Berelson, R. K. Anderson, O. Harkavy, J. Maier, W. P. Mauldin, & S. J. Siegel (Eds.), *Family Planning and Population Programs: A Review of World Developments*. Chicago: Univ. Chicago Press, 1966. Pp. 375-386.
17. SPEROFF, L., & RAMWELL, P. W. Prostaglandins in reproductive physiology. *Amer. J. Obstet. & Gynec.*, 1970, **107**, 1111-1130.
18. STYCOS, J. M. Some directions for research on fertility control. *Milbank Mem. Fund Quart.*, 1958, **36**, 126-148.
19. VON EULER, U. S. A depressor substance in the vascular gland. Proceedings of the Physiological Society, March 16, 1935. *J. Physiol.*, 1935, **84**, 21P-22P.
20. ———. On the specific vaso-dilating and plain muscle stimulating substances from accessory genital glands in man and certain animals (prostaglandins and vesiglandins). *J. Physiol.*, 1936, **88**, 213-234.
21. WESTOFF, C. F. The modernization of U.S. contraceptive practice. *Fam. Planning Persp.*, 1972, **4**(3), 9-12.
22. WESTOFF, C. F., & RYDER, N. B. Recent trends in attitudes toward fertility control and in the practice of contraception in the United States. In S. J. Behrman, L.

Corsa, Jr., & R. Freedman (Eds.), *Fertility and Family Planning*. Ann Arbor: Univ. Michigan Press, 1969. Pp. 388-412.

23. ZIEGLER, F. J. Vasectomy and adverse psychological reactions. *Ann. Intern. Med.*, 1970, **73**, 853.

*Institute of Personality Assessment and Research*

*University of California*

*2240 Piedmont Avenue*

*Berkeley, California 94720*

## PARENTAL DOMINANCE AND SEX-ROLE IDENTIFICATION IN SCHIZOPHRENIA\*

*George Williams College*

---

NICHOLAS D. KOKONIS

---

### SUMMARY

In the present study hypotheses derived from developmental-psychoanalytic and role theories of sex-role identification were tested in father-dominant and mother-dominant schizophrenic and normally functioning adult male Ss ( $n = 15$  in each group). Parental dominance and sex-role identification were measured by a battery of traditional paper-and-pencil tests. It was found that (a) father-dominant and mother-dominant schizophrenics scored significantly lower ( $p < .01$ ) than their respective normally functioning counterparts; (b) no significant differences existed between father-dominant and mother-dominant schizophrenics; and (c) irrespective of parental dominance, schizophrenics scored significantly lower ( $p < .01$ ) than normally functioning Ss in regard to sex-role identification. These findings were interpreted to support a psychodynamic theory of schizophrenia emphasizing the importance of dependence and security needs, while they failed to support role-theory interpretations of the process of sex-role identification.

---

### A. INTRODUCTION

Although many writers from Freud (7) to Arieti (1) and Searles (18) have postulated that inadequate sex-role identification is an important characteristic of schizophrenic patients, a review of the research literature (12) has produced inconclusive evidence. Some of these studies have turned up results that are difficult to compare because of methodological considerations, while others are exploratory or empirical, leaving much to be desired in terms of a theoretical point of view regarding the development of schizophrenia and of sex-role identification. On the assumption that schizophrenics do have an impaired sense of sex-role identification, how may this phenomenon be accounted for in a systematic, if not parsimonious, way? In the

---

\* Received in the Editorial Office on April 13, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

opinion of this writer, there are two possible theoretical views: (a) psychoanalytic-developmental theory making use of the concept of fixation and (b) role theory emphasizing inappropriate role models.

An interesting psychoanalytic-developmental theory of schizophrenia, and of psychopathology in general, is that of Roth, Berenbaum, and Hershenson (17). This theory, which they call the developmental model, views mental illness as a continuum along a scale of maturity, and maintains that schizophrenia is the psychopathology associated with fixation at the oral period of development (i.e., first year of life), of which the dominant needs are those of security and dependency. It would follow, then, that the schizophrenic's sex-role identification development is impaired inasmuch as sex-role identification is dependent upon normally going through and resolving the Oedipal conflict, as most psychodynamic writers have suggested (16). In regard to role-theory formulations of sex-role development, a review of the literature by this writer (12) has indicated that role theorists maintain that sex-role identification depends on the power of the identificand, a combination of his reward and his punishment potential (3). From this standpoint, it is important that the male child have a male model (i.e., father) who is the more dominant of the two parents in order that he can learn the masculine role appropriately, on the assumption that the home environment is of primary importance for such learning. Basic to this proposition is the finding that maternal dominance has a disruptive effect on the sex-role development of boys (9). It would follow, then, that only male schizophrenics from mother-dominant families have an impaired sense of sex-role identification.

The purpose of the present study was to investigate, within a theoretical frame of reference, the concept of sex-role identification in schizophrenia by means of a battery of reasonably dependable psychological tests. To this end, the following null hypotheses were formulated and subsequently tested:

*Hypothesis 1.* Schizophrenic Ss who come from father-dominant families do not differ significantly in their sex-role identification from schizophrenics who come from mother-dominant families.

*Hypothesis 2.* Schizophrenic Ss who come from father-dominant families do not differ significantly in their sex-role identification from normally functioning Ss who come from father-dominant families.

*Hypothesis 3.* Schizophrenic Ss who come from mother-dominant families do not differ significantly in their sex-role identification from normally functioning Ss who come from mother-dominant families.

*Hypothesis 4.* Schizophrenic Ss (father-dominant and mother-dominant subgroups combined) do not differ significantly in their sex-role identifica-

tion from normally functioning Ss (father-dominant and mother-dominant subgroups combined).

Confirmation of Hypothesis 1 would support psychoanalytic-developmental interpretations of sex-role identification, whereas rejection of this hypothesis would support role-theory interpretations. On the other hand, confirmation of Hypotheses 2 and 3 would support role-theory interpretations of the process of identification, whereas rejection of these hypotheses would support psychoanalytic-developmental interpretations. Finally, rejection of Hypothesis 4 would further support psychoanalytic-developmental theory, while confirmation of this hypothesis would not offer support for role theory, as this particular hypothesis is not directly related to role theory.

## B. METHOD

### 1. *Subjects*

Four groups, each consisting of 15 Ss were utilized in the present study: father-dominant schizophrenics, mother-dominant schizophrenics, father-dominant normally functioning Ss, and mother-dominant normally functioning Ss. All Ss were Caucasian, male, between the ages of 20 and 35, with from 12 to 16 years of education; none suffered from central nervous system disturbance or addiction of any kind. All Ss had both parents living.

Schizophrenics were selected on the basis of two independent and concurring diagnoses, whereas the normals were recruited from a variety of occupations and were interviewed prior to testing to control for history of emotional disturbance. The two groups did not differ significantly in statistical comparisons with respect to age, socioeconomic background, or marital status. Among schizophrenics, the mean cumulative length of hospitalization was two years for the father-dominant group and two and one-half years for the mother-dominant group. The schizophrenic sample included seven paranoid, six catatonic, eight acute undifferentiated, and nine chronic undifferentiated Ss that were evenly distributed throughout the schizophrenic subgroups.

All Ss participated voluntarily, and no compensation was offered for cooperating.

### 2. *Measures*

To measure parental dominance, Lu's (15) Dominance-Submission Scale was used as directed by its author. The scale was administered to both Ss



and their parents, and only those cases in which there was agreement between the two indices of dominance were included in the study.

Sex-role identification, defined as the internalization of the role considered appropriate to a given sex and the "unconscious reactions" characteristic of that role (4), was measured by Terman and Miles' (20) Word Association Test (WAT) and Emotional-Ethical Attitudes Test (EEA), Swensen's (19) Sexual Differentiation Scale (SDS), and Krout and Tabin's (14) Personal Preference Scale (PPS), Section VIII. As discussed elsewhere (12), the reliability and validity of these instruments are satisfactory for this type of study.

### 3. Procedure

The administration of the tests used was in accordance with the instructions given by their authors. To insure anonymity and more reliable data, the Ss were not requested to identify themselves on the test booklets. The total procedure required roughly two hours of time for normals and two and one-half hours for schizophrenics.

### C. RESULTS

An analysis of variance indicated no significant differences between father-dominant and mother-dominant schizophrenics on the measures of sex-role identification, thus supporting Hypothesis 1. However, the findings failed to support Hypothesis 2. As shown in Table 1, with the only exception of WAT, the father-dominant normally functioning Ss scored significantly higher ( $p < .01$ ) than father-dominant schizophrenics. Similarly, the results did not support Hypothesis 3, for, compared to mother-dominant schizophrenics, mother-dominant normally functioning Ss scored significantly higher ( $p < .01$ ) on all measures with the exception of WAT.<sup>1</sup> In the testing of Hypothesis 4, significant differences ( $p < .01$ ) were also noted between the schizophrenic group (father-dominant and mother-dominant Ss combined) and the normally functioning group (father-dominant and mother-dominant Ss combined). As presented in Table 1, these differences, rejective of Hypothesis 4, are all in favor of the normals. A two-way analysis

---

<sup>1</sup> The consistent failure of WAT to distinguish between groups is puzzling. Although no definitive conclusions have yet been reached, the reasons for the minimal discriminatory value of this measure would seem to lie in its relatively moderate reliability (.62) and, perhaps more importantly, in the changes in vocabulary and word usage observed in the general population over the past several years. In a personal communication to the author, Dr. Allan K. Rosenwald has suggested that this finding might also be related to the looseness of the schizophrenic's associational processes.

TABLE 1  
MEANS AND SDs ON MEASURES OF SEX-ROLE IDENTIFICATION FOR  
SCHIZOPHRENIC AND NORMAL FATHER-DOMINANT, MOTHER-DOMINANT,  
AND COMBINED GROUPS

Measure	Schizophrenics		Normals		D
	Mean	SD	Mean	SD	
<i>Father-Dominant Groups<sup>a</sup></i>					
Word Association Test (WAT)	-6.13	8.47	-2.80	8.20	
Emotional-Ethical Attitudes Test (EEA)	-2.33	22.61	40.07	21.39	42.40*
Sexual Differentiation Scale (SDS)	3.47	2.00	7.07	1.67	3.60*
Personal Preference Scale (PPS)	10.47	2.59	12.87	2.13	2.40*
<i>Mother-Dominant Groups<sup>a</sup></i>					
Word Association Test (WAT)	-6.07	9.61	-1.60	7.11	
Emotional-Ethical Attitudes Test (EEA)	-11.40	27.61	40.00	16.24	51.40*
Sexual Differentiation Scale (SDS)	4.27	2.15	8.27	.70	4.00*
Personal Preference Scale (PPS)	10.13	2.90	13.53	1.68	3.40*
<i>Combined Groups<sup>b</sup></i>					
Word Association Test (WAT)	-6.10	8.90	-2.20	7.57	
Emotional-Ethical Attitudes Test (EEA)	-6.87	25.22	40.03	18.66	46.90*
Sexual Differentiation Scale (SDS)	3.87	2.08	7.67	1.40	3.80*
Personal Preference Scale (PPS)	10.30	2.71	13.20	1.92	2.90*

<sup>a</sup> Cell entries are each based on 15 observations.

<sup>b</sup> Cell entries are each based on 30 observations.

\*  $p < .01$ .

of variance indicated no significant relationships or interaction effects between types of schizophrenia (i.e., paranoid, catatonic, acute undifferentiated, and chronic undifferentiated), parental dominance, and sex-role identification. Finally, no differences were noted between schizophrenics who were medicated and schizophrenics who did not receive medication. Table 2 summarizes the major findings of the present study.

#### D. DISCUSSION

The results of the present study clearly indicate that schizophrenics score significantly lower than normally functioning Ss on measures of sex-role identification. This statement applies equally for father-dominant and mother-dominant schizophrenics and is consistent with Roth *et al.*'s (17)

TABLE 2  
SUMMARY DATA INCLUDING HYPOTHESES TESTED, TESTS USED, AND THEORETICAL  
VIEW SUPPORTED BY THE FINDINGS OF THE SEX-ROLE IDENTIFICATION STUDY

Hypothesis	Test				Psychoanalytic versus role-theory			
	WAT	EEA	SDS	PPS	WAT	EEA	SDS	PPS
1	+	+	+	+	a	a	a	a
2	+	—	—	—	b	a	a	a
3	+	—	—	—	b	a	a	a
4	+	—	—	—	b	a	a	a

Note: WAT = Word Association Test; EEA = Emotional-Ethical Attitudes Test; SDS = Sexual Differentiation Scale; PPS = Personal Preference Scale; plus sign (+) = confirmation of the hypothesis tested by means of this test; minus sign (—) = rejection of the hypothesis tested by means of this test.

<sup>a</sup> Denotes evidence in support of psychoanalytic-developmental theory.

<sup>b</sup> Denotes evidence in support of role theory.

developmental model which states that, no matter how dissimilar in other respects (e.g., sex, age, symptomatology, etc.), schizophrenics are essentially alike in terms of developmental core issues: i.e., extreme dependence and security needs. The statement is also consonant with developmental-psychoanalytic interpretations of sex-role identification as being contingent upon the resolution of the Oedipal conflict around the fourth year of life. The human infant, it seems, must find ways of solving the tasks presented at a certain stage of growth before he can move on to the next stage with security. Evidently, the schizophrenic does not attain confidence in his sex-role identification, a sense of worth as a human being of his given sex, probably because of lack of security experiences and consequent insufficient sex-role identification. Johnson (10) has presented evidence to show that schizophrenics identify less with either parent as compared to normals, while Gardner (8) has reported that, as children, adult male schizophrenics exhibited marked sex-inappropriate behaviors. In the words of Searles (18), "the great problem of the preschizophrenic person, of course, is that, in keeping with the perpetuation, at the unconscious level, of the undifferentiated mother-infant stage of ego development, he has not achieved any deep-reaching sexual differentiation" (p. 92). This would seem to be the most parsimonious interpretation of the present findings, and possibly those of previous studies.

The present findings also suggest that the role of the father does not appear to be particularly relevant for the schizophrenic offspring's sex-role identification. The father-dominant schizophrenics in our study have no advantage over their mother-dominant counterparts as far as sex-role iden-

tification is concerned. These findings would seem to be inconsistent with role-theory interpretations of the process of sex-role identification, which stress the importance of parental power and dominance in identification (3). In light of the extensive evidence of conflict and pathology in the families of schizophrenics (2), it would seem reasonable to accept that "a hostile atmosphere is established in the home of the potential schizophrenic so that identification with either parent is precluded" (5, p. 354). However, lest it be inferred from the present data that the role of the father is not important at all in schizophrenia, it should be stressed that this study deals only with one facet of human development: namely, sex-role identification. The present findings do not negate the father's possible importance for the offspring's mastery of other kinds of developmental tasks.

One may still wonder why the present findings are so inconsistent with those of previous role-theory studies reviewed by the author (12), which indicated that in father-dominant families there is a greater tendency for boys to score higher on sex-role identification scales than boys in mother-dominant families. There are two important considerations. First, in previous investigations, parental dominance was measured in terms of *either* the Ss' or their parents' perceptions, whereas in the present study dominance was assessed by *both* Ss' and parents' perceptions. A second consideration is that nearly all previous studies utilized normally functioning Ss, while this study employed clinical (schizophrenic) Ss. In light of similar evidence pertaining to sex-role identification in neurosis (13), it is reasonable to conclude that parental dominance does not seem to be an important variable in sex-role identification studies of clinical groups. However, the developmental level reached by the S does seem to be a determining variable.

It should be stated, parenthetically, that no theoretical polarity is suggested in this paper between psychosexual and psychosocial determinants of behavior. The author recognizes that psychosexual development is one aspect of the psychosocial development of the individual, as lucidly described by Erikson (6) and the ego psychologists. For the sake of parsimony, these two viewpoints were treated in the present study in a way that might be misconstrued to imply two theoretically polar and mutually exclusive categories. However, no such polarity is implied.

Finally, the present findings, having implications for a psychodynamic interpretation of schizophrenia, do not necessarily conflict with interpretations of the disorder as a congenital or organically based condition, such as Kety *et al.*'s (11) theory of genetic transmission of vulnerability to schizo-



phrenia, for even these authors have concluded that their findings "also imply the requirement of nongenetic, environmental factors for the development of clinical schizophrenic illness" (p. 306; emphasis added).

## REFERENCES

1. ARIETI, S. Interpretation of Schizophrenia. New York: Brunner, 1955.
2. BOWEN, L. M. A family concept of schizophrenia. In D. Jackson (Ed.), *The Etiology of Schizophrenia*. New York: Basic Books, 1960. Pp. 346-372.
3. BRIM, O. G. Family structure and sex-role learning by children: A further analysis of Helen Koch's data. *Sociometry*, 1958, 21, 1-16.
4. BROWN, D. G., & LYNN, D. B. Human sexual development: An outline of components and concepts. *J. Mar. & Fam.*, 1966, 28(2), 155-162.
5. CAPUTO, D. V. The parents of schizophrenics. *Fam. Process*, 1963, 2, 339-356.
6. ERIKSON, E. H. *Childhood and Society*. New York: Norton, 1950.
7. FREUD, S. Introductory Lectures on Psychoanalysis. London: Allen & Unwin, 1922.
8. GARDNER, G. G. The relationship between childhood neurotic symptomatology and later schizophrenia in males and females. *J. Nerv. & Ment. Dis.*, 1967, 144, 97-100.
9. HETHERINGTON, E. M. A developmental study of the effects of sex of the dominant parent on sex-role preference, identification, and imitation in children. *J. Personal. & Soc. Psychol.*, 1965, 2, 188-194.
10. JOHNSON, M. H. Parental identification and schizophrenia. Unpublished dissertation, The University of Arizona, 1964. *Diss. Abst.*, 1964, 25(5), 3111.
11. KETY, S. S., ROSENTHAL, D., WENDER, PH H., & SCHULSINGER, F. Mental illness in the biological and adoptive families of adopted schizophrenics. *Amer. J. Psychiat.*, 1971, 128(3), 302-306.
12. KOKONIS, N. D. Sex-role identification in schizophrenia: Psychoanalytic and role-theory predictions compared. Unpublished Doctoral dissertation, Illinois Institute of Technology, 1971. *Diss. Abst.*, 1971, 32(3), 1849-1850.
13. ———. Sex-role identification in neurosis: Psychoanalytic-developmental and role-theory predictions compared. *J. Abn. Psychol.*, 1972, 80 (1), 52-57.
14. KROUT, M. H., & TABIN, J. Measuring personality in developmental terms: The Personal Preference Scale. *Genet. Psychol. Monog.*, 1954, 50, 289-335.
15. LU, Y. C. A study of the dominant, equalitarian, and submissive roles in marriage. Unpublished Doctoral dissertation, University of Chicago, Chicago, Illinois, 1950.
16. MUNROE, R. L. *Schools of Psychoanalytic Thought*. New York: Holt, 1955.
17. ROTH, R. M., BERENBAUM, H. L., & HERSHENSON, D. B. The developmental theory of psychotherapy: A systematic eclecticism. Unpublished paper, Illinois Institute of Technology, Chicago, 1967.
18. SEARLES, H. I. Sexual processes in schizophrenia. *Psychiatry*, 1961, 24, 87-95.
19. SWENSEN, C. H. Sexual differentiation on the Draw-A-Person test. *J. Clin. Psychol.*, 1955, 11, 37-41.
20. TERMAN, L. M., & MILES, C. C. *Sex and Personality: Studies in Masculinity and Femininity*. New York: McGraw-Hill, 1936.

George Williams College

555 Thirty-First Street

Downers Grove, Illinois 60515



## PSYCHOLOGICAL TEST PERFORMANCE AND INDIGENOUS CONCEPTIONS OF INTELLIGENCE\*<sup>1</sup>

*Institute of Nutrition of Central America and Panama and Brandeis University*

ROBERT E. KLEIN, HOWARD E. FREEMAN, AND RICARDO MILLETT<sup>2</sup>

### SUMMARY

The psychological test performance of 10 seven-year-old boys from an isolated, rural Guatemalan village was compared with ratings of their intelligence made by 42 adult village members. The adults' judgments were found to correlate strongly with the children's performance on the Embedded Figures, Verbal Analogies, and Memory-for-Designs Tests.

### A. INTRODUCTION

Cross-cultural studies of cognitive development frequently employ psychological tests originally developed and standardized on populations fundamentally different in cultural background from the subjects who are subsequently given the tests (1, 2, 3). Little is known about the congruence between scores on tests of intellectual abilities and estimates of intelligence by community members acquainted with the subjects.

The purpose of the present study was to compare, in an isolated Guatemalan village, preschool children's performance on tests of intellectual abilities with estimates of their "intelligence" by adults from their community. This exploratory study is part of a broad investigation of congruence between different social and psychological tests and indigenously judged social-role competence.

\* Recommended by Jerome Kagan of the Editorial Board, received in the Editorial Office on April 14, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The data for this study are drawn from the Guatemalan growth and development project by Contract #PH43-65-640 of the National Institute of Child Health and Human Development, N. I. H. This investigation is part of a program of collaborative research on Uniform Measures of Social Competence by H. E. Freeman, J. Kagan, R. E. Klein, and A. K. Romney and is supported by grant GS-33047 of the National Science Foundation.

<sup>2</sup> Requests for reprints should be sent to Robert E. Klein at the address shown at the end of this article. The second and third authors are at Brandeis University.

## B. METHOD

Systematic impressions of indigenous conceptions of intelligence were obtained in ethnographic interviews with adult members of an isolated, rural Guatemalan village. Considerable knowledge was available about this community of approximately 1000 Spanish-speaking inhabitants who barely subsist by the consumption and sale of agricultural production. This background information was accumulated as part of a long-term longitudinal study of nutrition and mental development taking place in this and three similar villages.

The most descriptive indigenous term for intelligence is *listura*. This word is most often translated back to English as "smartness." The behavioral characteristics used by adults in the village when describing children who are "*listo*" are "independence," "verbal facility," "good memory," "alterness," and a high level of physical activity.

Ten seven-year-old boys were selected randomly from the longitudinal study and photographed in a standing position against the same background. Randomly selected adults in this village were asked to name the boys in the photographs. This procedure was stopped after 42 judges were selected who knew all 10 boys. Each judge ranked the 10 boys on *listura*. The rankings were obtained by paired comparisons with use of the children's photographs.

A large number of psychological test scores were available on the subjects. Measures of perception, memory, and language were selected for this analysis. These tests will be discussed further in the Results section.

## C. RESULTS

There was a high degree of congruence among the ranking of the various judges. Agreement between any two judges averaged about 70%, and the relationship between judged ranks was significant at the .03 level. The rank-order matrix was subjected to multidimensional analyses, and the correlation between the resulting metric scale scores and the rank order was .95, suggesting that the ranks approximated the property of an interval scale. In Table 1, we show Pearsonian correlations between rank order scores and the psychological tests.

Although the values of the correlations are moderate to high, the sample size is small, and thus the confidence limits around the correlations are broad. Nevertheless, the results are provocative.

The *perception* measures are from an Embedded Figures Test. In this test, the child is scored on both how many correct responses he makes in locating figures embedded in complex pictures and the length of time between the

TABLE 1  
CORRELATIONS BETWEEN RANK ORDER SCORES AND THE PSYCHOLOGICAL TESTS

Measures	<i>r</i>	<i>p</i>
Perception: <i>Embedded Figures Test</i>		
Number of correct responses	.75	<i>p</i> < .02
Response time	.49	<i>p</i> > .10
Language:		
Verbal Analogies	.57	<i>p</i> < .10
Picture Vocabulary Test-Naming	.34	<i>p</i> > .10
Picture Vocabulary Test-Recognition	.11	<i>p</i> > .10
Memory: <i>Memory-For-Designs Test</i>		
Trial 1	.55	<i>p</i> < .10
Trial 2	.52	<i>p</i> > .10
Trial 3	.61	<i>p</i> < .10

item presentations and his responses. Judged *listura* is highly correlated with children's correct test responses, and is moderately related to their response times. It is of interest that there is also a positive correlation ( $r = .50$ ) between children's response time and number of correct responses on the Embedded Figures Test: children who have more correct responses take more time to answer. Thus, although adults emphasize high activity level in describing children who are "*listo*," these children are able to control their response rhythm and contemplate before answering when faced with a complicated task.

Of the three measures of *language*, one is moderately correlated with the judges' ratings. Responses on a Verbal Analogies Test are predictable from adult estimates of *listura*. However, there is a lower correlation between *listura* and the children's ability to name particular objects that appear in a set of pictures, and almost no relation between judge's estimates and recognition of the pictures that correspond with objects named by the examiner. The lack of correlation between judged *listura* and picture recognition is probably due to the relative ease of the task and the lack of variability in test scores.

The final three variables are scores for successive trials on the Memory-For-Designs Test. The designs are made up of four multicolored blocks. Children are given five seconds to study these designs and then must reconstruct them from memory. The correlations between *listura* and memory for designs are strong for all three trials.

#### D. DISCUSSION

The paired-comparison rating technique used here proved to be very useful. It was easily understood by the illiterate adult judges, and it provided very stable subject rankings.

However, since sample size is extremely small, the findings are only suggestive. Nevertheless, the popular conceptions of intelligence among adults in this rural Guatemalan village do show remarkable congruence with aspects of cognitive performance on tests adapted from batteries developed for use with urban children in the United States.

Expanded research efforts similar to this one may provide more systematic knowledge of the uniform characteristics of cognitive competence across cultures. Cultural variability and congruence in conceptions of cognitive development, along with the relationship of measures of cognitive development to social competence, are critical issues in the field of cross-cultural developmental psychology.

#### REFERENCES

1. CRAVIOTO, J., DELICARDIE, E. R., & BIRCH, H. G. Nutrition, growth and neuro-integrative development: An experimental and ecologic study. *Pediatrics*, 1966, **38**, 319-372.
2. JOHNSON, D. L., JOHNSON, C. A., & PRICE-WILLIAMS, D. The Draw-A-Man test and Raven Progressive Matrices performance of Guatemalan boys and Ladino children. *Rev. Interamer. Psicol.*, 1967, **1**, 143-157.
3. KLEIN, R. E., GILBERT, O., CANOSA, C., & DE LEÓN, R. Performance of malnourished in comparison with adequately nourished children on selected cognitive tasks (Guatemala). Paper presented at the Annual Meeting of the American Association for the Advancement of Science, Boston, Massachusetts, December, 1969.

*Division of Human Development*

*Institute of Nutrition of Central America and Panama (INCAP)*

*Apartado Postal 11-88*

*Guatemala City, Guatemala*

## SIGNATURE SIZE AND SELF-ESTIMATION: A BRIEF NOTE\*

*Department of Sociology/Anthropology, Western Washington State College*

---

E. R. MAHONEY

---

Zweigenhaft<sup>1</sup> recently published an article offering evidence for a relationship between signature size and social status. His findings were replicated by Swanson and Price<sup>2</sup> using a sample of school system personnel.

The hypothesis stated by Zweigenhaft was that signature size is a "projective test of that man's estimation of himself" (p. 49). The data used to test the hypothesis, however, did not bear directly on the hypothesized relationship. Conceptually, the hypothesis states that signature size is related to self-esteem, since this is a person's estimation of himself regardless of his or her social status.

In an attempt to test the originally stated hypothesis that signature size is related to a person's estimation of himself, the variable signature size was related to level of self-esteem. Subjects consisted of 227 undergraduate students enrolled in introductory sociology courses at a western state college.

Each subject completed a 15 line self-conception instrument (WAI) which requires the respondent to give 15 self-directed responses to the question: "Who Am I?" Self-esteem level is measured by the subject indicating his own positive-negative evaluation of those listed characteristics of himself using a five-point evaluation scale.<sup>3</sup> The mean of these evaluative ratings is defined as level of self-esteem. Each subject then responded with his signature on a plain 3 × 5 card after being told that the cards would not be used by the researcher, and thus signatures would not be connected with any responses. In order for the investigator to obtain signature size, subjects themselves placed the signed card under the last sheet of the questionnaire and indicated by vertical and horizontal line tracings the size of the signature as defined by Zweigenhaft. Subjects then indicated how many letters the signature contained and discarded the signed cards.

---

\* Received in the Editorial Office on April 14, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Zweigenhaft, R. L. Signature size: A key to status awareness. *J. Soc. Psychol.*, 1970, **81**, 49-54.

<sup>2</sup> Swanson, B. R., & Price, R. L. Signature size and status. *J. Soc. Psychol.*, 1972, **87**, 319.

<sup>3</sup> Gordon, C. Self-conceptions: Configurations of content. In C. Gordon & K. Gergen (Eds.), *The Self in Social Interaction*, New York: Wiley, 1968. Pp. 115-136.



Signature size was measured by taking the  $\text{cm}^2$  of the signature divided by the number of letters. This yielded a weighted  $\text{cm}^2$  signature size score for each subject.

To test the hypothesis that signature size is related to self-esteem, a Pearsonian  $r$  was computed between the two sets of above scores. The hypothesis was not supported ( $r = +.031, p > .05$ ).

Since approval motivation is consistently found to have a positive linear relationship to self-esteem, each subject also completed the Marlowe-Crowne Social-Desirability Scale<sup>4</sup> immediately after responding to the WAI instrument. In these data a moderate positive correlation was observed between social desirability scores and self-esteem scores ( $r = +.376, p < .001$ ). No significant relationship, however, was observed between signature size and social desirability ( $r = -.031, p > .05$ ). Thus, in the test of the hypothesis that self-esteem is related to signature size, when the effects of social desirability were partialled out, no support was found ( $r = +.022, p > .05$ ).

These data from a sample of undergraduate college students fail to support the stated Zweigenhaft hypothesis that signature size is related to a person's estimation of himself.

*Department of Sociology/Anthropology  
Western Washington State College  
Bellingham, Washington 98225*

---

<sup>4</sup> Crowne, D. P., & Marlowe, D. *The Approval Motive*. New York: Wiley, 1964.

## STATUS OF FRUSTRATOR AS A FACILITATOR OF AGGRESSION: A BRIEF NOTE\*

*Department of Speech, Bowling Green State University*

LAWRENCE J. CHASE AND NORBERT H. MILLS<sup>1</sup>

In 1968, Doob and Gross<sup>2</sup> demonstrated that the status of the frustrating agent is inversely related to the amount of aggression displayed by the frustrated individual. This was accomplished by varying the status of an automobile at several intersections in Palo Alto, California. Approximately half the trials conducted involved a high status auto, while the remainder utilized a low status auto. Observations were based upon the horn-honking responses of drivers blocked by the frustrator vehicle. It was found the low status auto elicited more honking than the high status car, and that males were more likely to honk regardless of frustrator status.

In attempting to replicate the above study, the present investigation examined the variables of status and sex. The conditions of the original experiment were followed. A 1972 Mercury served as the high status car; a 1968 Volkswagen, somewhat rusty and dull in finish, was used as the low status vehicle. In order to check the validity of the status manipulation, a questionnaire was administered to 30 subjects. The results showed that the 1972 Mercury was rated as being of significantly higher status than the 1968 Volkswagen ( $\chi^2 = 10.8$ ,  $df = 1$ ,  $p < .01$ ). Subjects for the field experiment were 40 drivers in the Toledo, Ohio, area: 20 in each status condition. A trial was counted if at least one vehicle pulled up behind the experimental car and came to a complete stop. Observers recorded the latency of each honk and honking frequency. The experimental car terminated blockage only after two honks had sounded or 15 seconds had elapsed, whichever occurred first.

Subjects honked more readily at the high status car, contrary to the findings of Doob and Gross. In the high status condition, the mean latency of the first honk was 6.25 seconds, in contrast to a mean latency of 10.45 seconds noted for drivers in the low status condition ( $t = 2.36$ ,  $df = 38$ ,  $p < .05$ ). No significant difference was found between male and female drivers; 73%

\* Received in the Editorial Office on April 16, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The authors wish to express their gratitude to Dr. Raymond K. Tucker and Mr. James E. Sayer for their invaluable assistance in this effort.

<sup>2</sup> Doob, A. N., & Gross, A. E. Status of frustrator as an inhibitor of horn-honking responses. *J. Soc. Psychol.*, 1968, 76, 213-218.

of the males honked at least once, as compared to 69% of the females subjects ( $\chi^2 = .10$ ,  $df = 1$ ,  $p > .05$ ). This finding also fails to support the results of the original experiment.

The differences between the results obtained in the previous study and the present study may be due to shifting mores with the passing of time (1967 vs. 1972) or to regional variation (California vs. Ohio).

*Department of Speech*  
*Bowling Green State University*  
*Bowling Green, Ohio 43403*

## A POLITICAL DIMENSION IN THE I-E SCALE: A BRIEF NOTE\* 1

*University of Rochester*

---

MIRON ZUCKERMAN

---

Two recent factor analyses<sup>2,3</sup> of Rotter's I-E scale identified two factors: a general one and a political one that focuses on the belief in citizens' control over state and world affairs. These findings suggest that the I-E scale may be divided into subareas of different content. The present study examines the differential impacts of the recent presidential election on the response to the political (items 3, 12, 17, 22, and 29) and the nonpolitical (all other items) parts of the I-E scale.

In the period preceding the election, public polls showed Nixon as the probable winner in spite of an extensive Democratic campaign. This should lead to a greater "externality" among McGovern's supporters than among Nixon's supporters; i.e., McGovern's supporters should score higher on the I-E scale. In addition, those who actually voted for McGovern were obviously disappointed with the results, whereas those who voted for Nixon achieved their desired goal. This should further increase the externality of the former and increase the internality of the latter. It was hypothesized that the differences between McGovern's and Nixon's supporters before the election and the increase or decrease in externality after the election should be more pronounced on the political than on the nonpolitical part of the I-E scale.

Subjects were 86 University of Rochester undergraduates who were eligible to vote. They filled out the I-E scale two weeks before the election and then again one day after the election. On the second occasion they were asked to fill out the questionnaire as if for the first time. In the second session subjects were also asked to state for whom they had voted or, if they had not voted, for whom they would have voted. Subjects ( $N = 6$ ) who voted for a candidate other than McGovern or Nixon or who could choose between the two were dropped from analysis. The rest were divided into (a) McGovern's sup-

---

\* Received in the Editorial Office on April 16, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The author would like to thank Marsha S. Lipets for her assistance in this investigation.

<sup>2</sup> Mirels, H. L. Dimensions of internal versus external control. *J. Consult. & Clin. Psychol.*, 1970, 34, 226-228.

<sup>3</sup> Zuckerman, M., & Gerbasi, K. Dimensions in the scale. Unpublished manuscript, University of Rochester, Rochester, New York, 1973.

porters—i.e., those who voted ( $N = 56$ ) or would have voted ( $N = 2$ ) for McGovern—and (b) Nixon's supporters—i.e., those who voted ( $N = 16$ ) or would have voted ( $N = 6$ ) for Nixon.

Before the election McGovern's supporters were found to be significantly more external on the political subscale than Nixon's supporters ( $M = 2.7$ ,  $M = 1.8$ , respectively;  $t = 2.02$ ,  $df = 78$ ,  $p < .05$ , two tail). However, the difference between McGovern's and Nixon's supporters on the remaining part of the I-E scale was not significant ( $M = 10.4$ ,  $M = 9.7$ , respectively;  $t = .67$ ,  $df = 78$ ). In addition, the correlation between the political and the nonpolitical part of the I-E scale was .76 ( $p < .001$ ) among Nixon's supporters, but only .26 ( $p = .05$ ) among McGovern's supporters. The difference between the two correlations ( $p < .02$ ) further supports the hypothesis that the pre-election situation affected the political externality of McGovern's supporters, thus reducing its correlation with the remaining part of the I-E scale.

The comparison between the pre- and postelection mean scores on the political and nonpolitical parts of the I-E scale did not show any significant difference. Those who voted for McGovern increased their externality by .23 on the political and by .20 on the nonpolitical parts of the I-E scale ( $t = .28$ ,  $df = 55$ ;  $t = .20$ ,  $df = 55$ , respectively). Those who voted for Nixon also increased their externality by .30 on the political and by .34 on the nonpolitical parts of the I-E scale ( $t = .28$ ,  $df = 15$ ;  $t = .19$ ,  $df = 15$ , respectively). The failure of the voting to change the I-E scores may have been due to the fact that the results were completely expected and failed to have an impact on the subjects' perception of their political control.

*Department of Psychology*  
*University of Rochester*  
*River Campus Station*  
*Rochester, New York 14627*



## THE DEVELOPMENT OF PERSONAL SPACE SCHEMATA TOWARD BODY BUILD\*<sup>1</sup>

*Eastern Michigan University*

---

RICHARD M. LERNER<sup>2</sup>

---

### SUMMARY

The development of personal space schemata toward Endomorph, Mesomorph, and Ectomorph body builds was investigated in a study of male and female kindergarten through third grade children. Subjects were instructed to move a marker along a board to indicate how close they wanted to come to stimulus figures representing each of the three body types. Results indicated that females used greater spatial distances than did males at all age levels, and that in grades one through three greater distance was used toward the Endomorph than towards either the Mesomorph or Ectomorph. Findings were discussed in relation to previous studies of personal space and of body build stereotypes.

---

### A. INTRODUCTION

In a series of studies (1, 2, 3, 4, 5, 6, 7, 8, 13) the attitudes that people of various ages hold towards different body builds have been assessed. All studies report that positive attitudes are maintained toward those having an Average, or Mesomorphic, physique, while negative stereotypes are held towards Chubby, or Endomorphic, and Thin, or Ectomorphic, body types. These stereotypes have been found to be related to the inculcation of positive self-concepts among Average build children and negative self-concepts among Chubby children (5).

To date, however, the implications of these stereotypes for the development of interpersonal relations involving children with different physique types have been relatively unexplored. What is the relation between the attitudes that people hold toward others having fat, thin, or average body types and

---

\* Received in the Editorial Office on April 20, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was supported in part by National Science Foundation Grants Numbers 25651-42 and 25652-26. The author wishes to thank Stuart A. Karabenick, Murray Meisels, and Henry Orloff for critical readings of the manuscript.

<sup>2</sup> Reprint requests should be sent to Richard M. Lerner at the address shown at the end of this article.

the behavior shown toward these physique groups? To begin exploration of this general question the present study investigated one aspect of interpersonal relationships, the development of personal space schemata toward body build.

Little (9, p. 237) has defined personal space as "the area immediately surrounding the individual in which the majority of his interactions with others take place." Little (9) and Meisels and Guardo (11) have operationalized personal space in terms of using a projective device whereby subjects manipulate the distance between human figure drawings or statuettes. Little (9) found that the distance placed between members of dyads decreased as the degree of attributed "liking" increased. This inverse relation between use of personal space towards a person and degree of psychological or interpersonal closeness has also been obtained with respect to such dimensions as friendly-unfriendly, handicapped-not handicapped, and deviant-not deviant (10, 11).

Accordingly, it was predicted that if the attitudes that children maintain towards fat, thin, and average physique types do correspond to some dimensions of their interpersonal behavior, then with a projective index of personal space similar to that of Meisels and Guardo (11), greatest personal space distance should be seen towards the fat person, next to the thin person, and the least in respect to the average build stimulus person. Because the average build person has favorable attitudes maintained towards him, and possesses a physique that may be considered not deviant and not handicapped, the least amount of personal space should be maintained toward him. Conversely, because the fat person (and to some extent the thin person) is afforded negative attitudes and has a body build that may be considered deviant, or handicapped (12), greater personal space should be used. These predictions were tested in a study of four groups of children, kindergarten, first, second, and third grade boys and girls.

## B. METHOD

### 1. Subjects

Eighty-six children from an elementary school in a semirural section of southeast Michigan were subjects. There were 20 kindergarten children (12 males, eight females; mean age = 5.8 years,  $SD = .57$ ), 22 first grade children (seven males, 15 females; mean age = 6.9 years,  $SD = .27$ ), 19 second grade children (11 males, eight females; mean age = 8.1 years,  $SD = .47$ ), and 25 third grade children (14 males, 11 females; mean age = 9.3 years,  $SD = .69$ ).

## 2. *Materials*

The projective index of interpersonal spatial distance was measured by use of a 61 cm wide  $\times$  45 cm high board covered with green felt. A 30 cm long line was cut through the board 15 cm from the lower edge, centrally spaced from either side. Along this line a 22.5 cm red marker (1.25 cm in width) could be moved. The handle of the marker extended from the base of the board and allowed the subject to grasp the handle completely in order to move the marker and indicate the preferred distance from selected stimulus figures that were placed at the left end of the line. Seven 22.5 cm high figure drawings were constructed by a professional artist. Four of these drawings (of a tree, a teddy bear, an ice cream cone, and a syringe) had not been used in previous research; three body build stimulus figures—of a chubby, average, and thin boy, respectively—had been used by Lerner and Korn (5).<sup>3</sup>

## 3. *Procedure*

All subjects were tested individually, either in an area immediately outside their classroom or within their classroom but separated from the rest of the class by about 20 feet and a partition. All children present in a class on the day of testing were assessed.

*a. Training.* First, the experimenter placed a picture of a tree on the left hand side of the center line. The experimenter then told the subject that they were going to play a game called "Coming close to things" and demonstrated that the subject could move the marker any place along the line in order to indicate how close he wanted to come to the picture. The experimenter then allowed the subject to practice this task with the tree picture and with another picture, of a teddy bear. All subjects were able to perform this task.

*b. Test.* The experimenter randomly and successively presented a side-view figure drawing of a male Endomorph, and Mesomorph, and an Ectomorph. With each presentation the subject was told, "Here is a picture of a boy your age that I know." The subject was asked to show the experimenter how close he would like to come to the boy by moving the marker. All trials began with the marker placed at the far end of the line. After the subject had responded to all three pictures, the stimuli were presented again, in a different random order, so that response reliability could be assessed.

*c. Posttest.* Although subjects might have had differing personal space schemata towards the three body types, it was possible that the measure of these schemata in this study might not have reflected these differences. To

<sup>3</sup> Copies of the stimulus figures can be obtained from the author upon request.

test this, subjects were asked to give preferred distances from objects having a high probability of evoking different affective reactions. These stimuli consisted of figure drawings of an ice cream cone and of a syringe and were assumed, *a priori*, to evoke positive and negative affect, respectively, and thus result in different personal space usage. These figures were presented twice to each subject in counterbalanced order.<sup>4</sup>

## C. RESULTS

### 1. Response Reliability

Test-retest response reliability for the kindergarten subjects was rather low, averaging about  $+ .4$  across the three body build stimuli. For the older groups, however, response reliability was moderately high: i.e., about  $+ .6$ ,  $+ .7$ , and  $+ .7$  for the first, second, and third graders, respectively. Similarly, reliability of the posttest measures increased from an average across the two posttest stimuli of  $+ .4$  for the kindergarten subjects, to  $+ .7$  for each of the older groups. Although in previous research the reliability of projective personal space responses has not been standardly assessed, these findings suggest that such responses become increasingly more stable and reach acceptable levels of reliability at relatively young ages.

### 2. Personal Space Towards Body Build

The data for the personal space distances were analyzed by means of a  $4 \times 2 \times 3$  analysis of variance, with dimensions consisting of Grade, Sex, and Body Build (within-subjects), respectively. An unweighted means analysis was used because of the presence of unequal cell *N*s (14). The results of this analysis, for both sets of personal space measures, are in part consistent with our predictions. For the initial set of three personal space measures—towards the Endomorph, the Mesomorph, and the Ectomorph—a significant main effect for Grade ( $F = 3.2$ ,  $df = 3, 78$ ,  $p < .05$ ), Sex ( $F = 8.4$ ,  $df = 1, 78$ ,  $p < .01$ ), and Body Build ( $F = 11.7$ ,  $df = 2, 156$ ,  $p < .01$ ) obtained; in addition there was a significant Grade  $\times$  Body Build interaction ( $F = 2.5$ ,  $df = 6, 156$ ,  $p < .05$ ).

In respect to the main effect for grade, third graders used significantly more space than the other groups (an average of 8.5 cm, 6.8 cm, 8.3 cm, and 12.0 cm for the kindergarten through third graders, respectively). This greatest use of space by the oldest subjects is a finding inconsistent with some of the

<sup>4</sup> The author thanks Stuart A. Karabenick for suggesting the inclusion of this posttest.



results of Meisels and Guardo (11), who reported that children's spatial schemata generally change with age in the direction of decreased spatial usage. The main effect for Sex, which was attributable to the fact that females consistently used more space (mean spatial distance for females = 10.7 cm) than males (mean distance for males = 7.1 cm) is, however, consistent with some of the results of Meisels and Guardo (11). These authors found that there was a strong tendency for females to maintain greater spatial distance than males in neutral or negative affect situations involving males.

The main effect for Body Build was accounted for by the fact that subjects used more space toward the Endomorph figure than toward either the Mesomorph or the Ectomorph; spatial usage towards these latter two figures did *not* differ. Thus, the average distance used towards the Endomorph was 11.5 cm, while this value was 8.2 and 7.1 for the Mesomorph and Ectomorph, respectively. Although the finding that greater spatial usage obtained toward the Endomorph confirms our prediction, the nonsignificant difference in spatial distances towards the Mesomorph and Ectomorph is inconsistent with our hypothesis. Finally, the significant Grade  $\times$  Body Build interaction was attributable to the fact that at the kindergarten level spatial usage was not significantly related to the body build stimuli presented, while the responses of the older groups indicated more spatial usage towards the Endomorph than towards the other two stimuli.

An analysis of variance computed for the second set of personal space measures replicated those of the previous one, with the exception that the main effect for Grade was not significant, although the means were in the same direction as in the first set of measures. This failure to replicate our previous result, which was inconsistent with the findings of Meisels and Guardo (11), may be attributable to several factors (e.g., the reliability level of this projective response, subject fatigue) but, in any event, these results suggest that the developmental course of spatial usage needs further investigation.

### 3. Posttest Measures

Although differences in personal space usage towards the body builds did obtain, it is of interest to note briefly the results of the posttest measures. An analysis of variance (for unequal *N*s) was used for each set of posttest measures (i.e., for each syringe-cone presentation). Our *a priori* reasoning, that greater spatial usage towards the syringe should obtain if our technique provided a sensitive index of spatial schemata, was confirmed: subjects used significantly more space with the syringe, mean distance = 12.9 cm, than with



the ice cream cone, mean distance = 1.8 cm ( $F = 119.8$ ,  $df = 1, 78$ ,  $p < .01$ ). These results were replicated in the analysis of the second set of post-test measures.

#### D. DISCUSSION

It has been well documented that children of the ages assessed in this study consistently maintain negative stereotypes to Endomorphs and Ectomorphs, and positive stereotypes to Mesomorphs (e.g., 3, 5, 13). The results of this study suggest only a partial correspondence between these attitudes and children's projective use of personal space. Consistent with a negative stereotype toward chubbiness, greatest space was used toward the Endomorph. However, no difference in spatial usage obtained between the similarly negatively stereotyped Ectomorph and the positively evaluated Mesomorph.

Several possible interpretations of this finding may be offered. One possibility is that subjects did not discriminate between the average and thin body types. The tenability of this interpretation is limited, however, because of a series of findings (e.g., 4, 5, 8) that indicate that similarly aged children can reliably discriminate such body types. Another possibility is that a developmental "lag" exists between the time at which body build stereotypes are acquired and the time at which they are isomorphically represented through personal space usage. Of course, this interpretation of the results is tenable only to the extent that the projective measure used here predicts behavioral spatial usage. While Little (9) has assumed such measures to be isomorphic with behavioral spatial usage, Meisels and Guardo (11) point out that whether this is the case is still an open issue. The position taken here is that the precise correspondence between projective and behavioral measures of personal space still requires direct empirical determination.

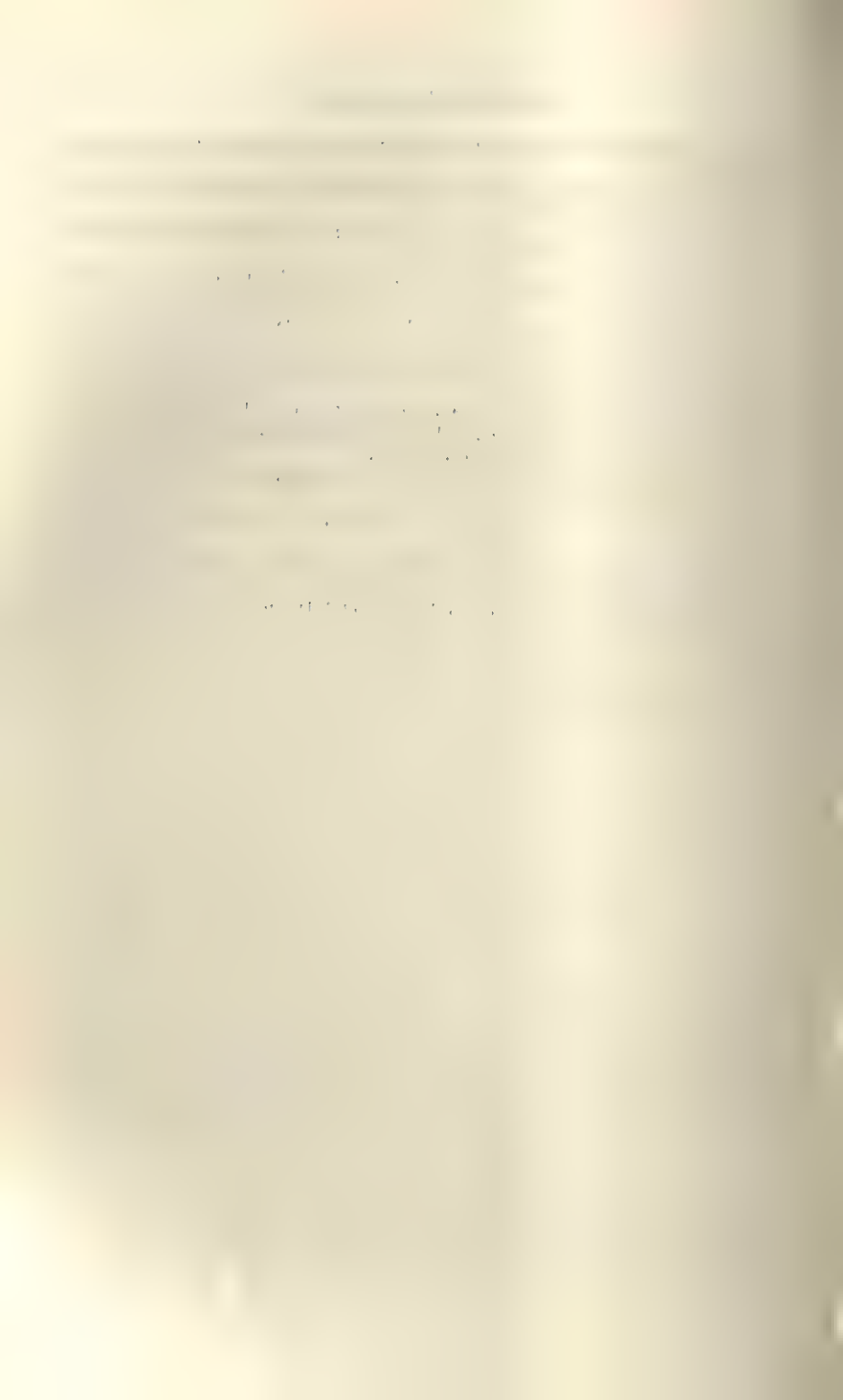
It seems clear then that future research should assess the development of personal space usage towards body build through the use of interpersonal behavioral measures. Previous research (e.g., 5) has provided considerable evidence indicating that the personality development of chubby children, as compared with average build children, is quite unfavorable. If future research extends the present findings to the domain of actual interpersonal spatial usage, then this would suggest that the socializing environment of chubby children would similarly portend unfavorable social development.

#### REFERENCES

1. LERNER, R. M. The development of stereotyped expectancies of body build-behavior relations. *Child Devel.*, 1969, **40**, 137-141.
2. ———. Some female stereotypes of male body build-behavior relations. *Percept. & Motor Skills*, 1969, **28**, 363-366.

3. ———. "Richness" analyses of body build stereotype development. *Devel. Psychol.*, 1972, 7, 219.
4. LERNER, R. M., & GELLERT, E. Body build identification, preference, and aversion in children. *Devel. Psychol.*, 1969, 1, 456-462.
5. LERNER, R. M., & KORN, S. J. The development of body build stereotypes in males. *Child Devel.*, 1972, 43, 912-920.
6. LERNER, R. M., & POOL, K. B. Body build stereotypes: A cross-cultural comparison. *Psychol. Rep.*, 1972, 31, 527-532.
7. LERNER, R. M., & SCHRODER, C. Kindergarten children's active vocabulary about body build. *Devel. Psychol.*, 1971, 5, 179.
8. ———. Physique identification, preference, and aversion in kindergarten children. *Devel. Psychol.*, 1971, 5, 538.
9. LITTLE, K. B. Personal space. *J. Exper. Soc. Psychol.*, 1965, 1, 237-247.
10. MEISELS, M., & CANTER, F. M. Personal space and personality characteristics: A non-confirmation. *Psychol. Rep.*, 1970, 27, 287-290.
11. MEISELS, M., & GUARDO, C. J. Development of personal space schemata. *Child Devel.*, 1969, 40, 1167-1178.
12. RICHARDSON, S. A. Handicap, appearance and stigma. *Soc. Sci. & Med.*, 1971, 5, 621-628.
13. STAFFIERI, J. R. Body build and behavioral expectancies in young females. *Devel. Psychol.*, 1972, 6, 125-127.
14. WINER, B. J. *Statistical Principles in Experimental Design* (2nd ed.). New York: McGraw-Hill, 1971.

Department of Psychology  
Eastern Michigan University  
Ypsilanti, Michigan 48197



## PERCEPTION OF ENVIRONMENTAL MODIFIABILITY AND INVOLVEMENT IN ANTIPOLLUTION ACTIVITIES\*<sup>1</sup>

*Department of Psychology, Texas A&M University*

---

HANNA LEVENSON

---

### SUMMARY

Members of an antipollution group were significantly less optimistic about the possibilities of cleaning up the environment than Nonmembers. Members expressed views that indicated that they were less in favor of technological means to lessen pollution and believed restoration of the environment would be difficult.

---

### A. INTRODUCTION

Recently there has been much talk about pollution, waste, conservation, and environmental quality. The present study was an attempt to account for some of the correlates of citizen involvement in voluntary, antipollution activities. Since it was reasoned that a person would not expend much effort in areas considered hopeless, it was hypothesized that people active in antipollution activities would perceive to a greater extent that something positive could be done to clean up the environment than those not so involved.

### B. METHOD

The subjects were 96 male and female adults in a southwestern metropolitan area. One-third of the subjects were selected randomly from the membership list of a local antipollution group (Members). Another third were selected from those people who had been sent a letter of notification of the group but did not join, and the final third had not been notified of the group (Nonmembers). Subjects were matched for sex and socioeconomic status (occupation and residence).

As part of a larger study, a series of scales were administered to each subject during an individual interview session. Included with the measures was

---

\* Received in the Editorial Office on April 23, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research formed part of a dissertation submitted to the faculty of Claremont Graduate School, Claremont, California, and was supported by a Woodrow Wilson Dissertation Fellowship.

the Environmental Modifiability Test (EMT). The EMT constructed for this study was patterned after items loading on a system modifiability factor found in a study dealing with the motivational dynamics of Negro youth (2). The system modifiability items were reworded for the EMT so as to apply to the area of environmental pollution. The EMT is composed of six items, half worded negatively and half positively, in a Likert format. A high score indicates that the subject feels that pollution can be lessened. Pretesting indicated that the EMT was not significantly correlated with the Marlowe-Crowne Social Desirability Scale (1) and had high internal consistency.

### C. RESULTS

Contrary to expectation, results indicated that Members scored significantly lower on the EMT than did Nonmembers ( $F = 4.5$ ,  $df = 1/88$ ,  $p < .05$ ). An examination of the content of the items helped to explain why Nonmembers would be more optimistic about the possibilities of cleaning up the environment. The six items on the EMT deal with the powers of society and technology to eliminate or lessen pollution, indicating a belief in the dominance of man over nature (3). The attitudes of the Members expressed during the interview appeared to be quite different, emphasizing the desire to live in harmony with nature and to lessen technological interference. Perhaps the Members' lower scores on the EMT, then, is not an expression of a "there's no use" attitude, but rather one of abhorrence at additional technological innovations.

Additionally, the items on the EMT tend to stress the ease with which the environment can be restored. Members, however, did not appear this optimistic—stating that a completely new life style and different economic and moral values might have to be adopted before pollution could be lessened.

With data analysis limited to correlational measurement, results are only suggestive. However, one may speculate that the numerous antilitter campaigns emphasizing the ease with which a clean environment can be obtained may result in obverse effects, with people becoming less concerned and less involved in antipollution efforts. More extensive research relating perception of the environment to antipollution involvement is clearly warranted.

### REFERENCES

1. CROWNE, D. P. & MARLOWE, D. A new scale of social desirability independent of psychopathology. *J. Consult. Psychol.*, 1960, **24**, 349-354.
2. GURIN, P., GURIN, G., LAO, R. C., & BEATTIE, M. Internal-external control in the motivational dynamics of Negro youth. *J. Soc. Issues*, 1969, **25**, 29-53.



3. KLUCKHOHN, F. R., & STRODTBECK, F. L. Variations in Value Orientations. Evanston, Ill.: Row, Peterson, 1961.

*Department of Psychology*  
*Texas A&M University*  
*College Station, Texas 77843*



## DEVELOPMENTAL TRENDS IN WISHES, CONFIDENCE, AND THE SENSE OF PERSONAL CONTROL FROM CHILDHOOD TO MIDDLE MATURITY\*<sup>1</sup>

*The Ohio State University and Columbus (Ohio) Public Schools*

---

JOHN E. HORROCKS AND MILTON C. MUSSMAN

---

### SUMMARY

The three wishes of a random sample of 100 subjects were obtained from each of five age groups: elementary, junior high, high school, early college (18-24 years), and late college (25-50 years). The wishes were placed in appropriate categories and into subcategories including possessions, activities, maintenance, achievement, and altruism. Exploratory measures of "confidence" and "sense of personal control" were taken for each subject. Sex and socioeconomic condition were analyzed. Results supported previously established trends of increasing generality, increasing altruism, and decreasing materialistic content up to different ages in adolescence and early adulthood depending on the data chosen to relate to the generalizations. In addition, a gradual increase in achievement wishes through middle adulthood was observed. The measure of "sense of personal control" was found to yield significant variation, while the data related to "confidence" varied only within chance limits. Comparisons with previous studies led to speculation that achievement orientations of adolescents may have lessened over the past decade.

---

### A. INTRODUCTION

The range and nature of an individual's wishes are useful sources of information in considering his level of functioning. For many years, wishes have been considered to provide information relative to individual functioning. Clinicians and others find in wishes representations of needs and motives, as well as indications of goals, interests, attitudes, beliefs, and values. An

---

\* Received in the Editorial Office on April 25, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press

<sup>1</sup> The authors are indebted to the administrators, teachers, and students in the Columbus (Ohio) Public Schools who generously contributed their time in generating data used in this study

early large scale study of children's wishes was reported by Washburne (8) who found that wishes do contain a degree of unconscious projection and that socially maladjusted populations have greater asocial, conflicting, and hostile wishes. Jersild (7) reported that more intelligent children of the same age tend to have wishes that are more inclusive and less materialistic.

Normative trends have been verified over the past 40 years. From childhood to early maturity it has been found that wishes become more inclusive, more altruistic, and less materialistic. Sex differences have been established into middle adolescence, with girls having more social and altruistic wishes, and boys being more concerned with personal achievement, wealth, and general self benefits (1, 3, 7, 8, 9). Socioeconomic variations were found by Horrocks and Flory (6) in that lower socioeconomic class adolescents wished more for wealth, marriage, and education, while the higher socioeconomic group showed more concern for improving humanity, having good health, and being happy. The wishes of the lower socioeconomic group were observed to show a striking similarity to the pattern reported by Washburne obtained during the depression 30 years earlier, with wealth, education, and an automobile being common to both groups in the five most frequent wishes.

The manner in which wishes are solicited from subjects has been found to have a significant influence on the nature of wishes obtained. Washburne asked first for three principal wishes, second for all the real wishes that could be thought of in three minutes, and finally for disapproved-of wishes. Boynton (2) and Gray (4) asked for only one wish. Cobb (3) varied his approach systematically, finding that the manner of the approach significantly affected the pattern of wishes and, further, that the open-ended stem, "I wish more than anything that . . .," resulted in a greater diversity of responses.

The present study attempts to provide current data on wishes, not only for comparison with previous findings, but also to extend wish data into middle adulthood. The method of investigation lent itself easily to obtaining exploratory information relative to trends in several personality variables which have been studied experimentally. One set of data presented in the present study resulted from asking the respondents to rate their ranked wishes in terms of their expectation that the wish would come true. It was hypothesized that the rating of a highly ranked wish as coming true would indicate the presence of a higher degree of confidence as a personality characteristic. Another set of data was generated by asking respondents to indicate which of their wishes could be influenced in outcome by their own personal effort. A high number of wishes so indicated might be assumed to reflect a high sense of personal control or internal locus of control.

## B. METHOD

During the winter of 1971, teachers in four elementary schools, two junior high schools, and one senior high school in a large American midwestern city were asked to solicit three wishes from each of their pupils in an anonymous, standardized fashion, using the questions, "If you could have any three wishes come true, what would they be?" After a short pause, they were to ask their pupils to list their wishes in rank order. Pupils were then asked to place plus and minus signs in front of the wishes that had the best and least chance of coming true and, further, to place stars in front of the wishes they felt they could do something about through their own efforts. During the following summer, the same procedure was used in two large undergraduate-graduate lecture courses at a university in the same city. From the large number of responses obtained, 100 sets of three wishes (50 male, and 50 female) were randomly selected to represent each group: the elementary (grades 3-6), junior high, senior high, early college (18-24 years), and late college (25-50 years). Two of the four elementary schools and one of the two junior highs were used to represent the low socioeconomic variable only and were not included in the age-trend analysis which used predominantly middle class populations.

Preliminary to setting up categories for recording the wishes, a review of categories used in previous studies was undertaken. It was found that considerable variation in the naming of categories existed. Little information was found to aid in the translation of wishes into categories that would increase comparability of results across the various studies. In the present study, therefore, an attempt was made to set up broad categories similar to those named in other research, with theoretical usefulness as an added criterion. A sampling of wishes was then used to establish the smaller categories within the broad areas already created. The full sample of wishes was then individually interpreted to establish the frequencies that became the data base for analysis. For the purposes of this study all wishes were given equal weight regardless of ranking.

The only statistical tool used to determine significance was the standard error of the difference between percents. While it was possible to determine expected frequencies in the sex variable and hence permit use of chi-square analysis, most of the other data did not permit clear prediction of expected frequencies. For this reason, and for reasons of clarity and comparability, it was felt that the results should be reported and analyzed in terms of percents and standard errors alone.



## C. RESULTS

The three wishes of each respondent were placed into categories and areas of categories as follows. *Possessions*: wealth, motor vehicles, bicycle, pets, clothes, miscellaneous. *Activities*: sports-games-music-etc. *Maintenance*: happiness-satisfaction, health-long life, internal states, freedom, appearance-personality, peer relations, love-have lover, marriage, family relationships, living conditions, good job, miscellaneous. *Achievement*: success-recognition, ability, school-education, vocation. *Altruistic*: world peace, improve humanity, welfare of loved one. *Miscellaneous*: more wishes, power, avoidance-school, avoidance-other.

Two principles guided the categorization process. It was felt that interpretation of the results would be enhanced if minor categories could be combined wherever theoretically possible. Thus the category *pets* included wishes for a horse, *health-long life* was felt to safely represent wishes for both conditions, and *improve humanity* was developed to represent wishes regarding pollution, helping deprived populations, and reducing prejudice. It was also felt that evidence in relation to previously established age trends could be more clearly presented if categories could be placed in areas that should reflect those trends from age to age.

Figure 1 presents the trends in percents of wishes for the five major areas. Wishes for possessions are seen to decrease from the elementary years to the 18-24 year group and then increase during the middle years, with the category *wealth* accounting for the change as may be observed in Table 1. Since wealth for an older adult may not clearly be equated with the ability to purchase possessions, it is felt that the previously established age trend towards increasing generality and less materialistic content of wishes is supported through early maturity and at least not refuted in the older group. The decreasing interest in activities after junior high may be observed and may at least be partially accounted for by the principle of increasing generality with age, although undoubtedly activities do not provide the same level of satisfaction for adults as they do for children and adolescents. Wishes for maintenance, achievement, and altruistic experiences may in general be seen to grow with age, although some exceptions may be observed. Maintenance categories involving wishes for adequacy, security, acceptance, problem reduction, etc. are of considerably less concern to the older group in comparison to the younger adults, with the categories *love-have lover* and *marriage* showing the greatest percent decrease. Wishes for achievement were relatively low in junior high school compared to the levels of the elementary and senior

high groups. Altruism in senior high is lower in frequency than either in junior high or the 18-24 year college group, the difference being found in the *improve humanity* category. The difference between the senior high and the college group is significant in the area of altruism. These findings have some implication for age trends from early through late adolescence. The trends shown in the general areas thus support earlier data on increasing generality, increasing altruism, and decreasing materialism with age, with some age-specific variations. The trend from junior high through middle adulthood for achievement wishes to increase continuously over a four standard error span is felt to add significant information to the hypothesis dealing with growth of achievement motivation with age. This finding is not clearly supported in data from past research and may be a function of the present method which

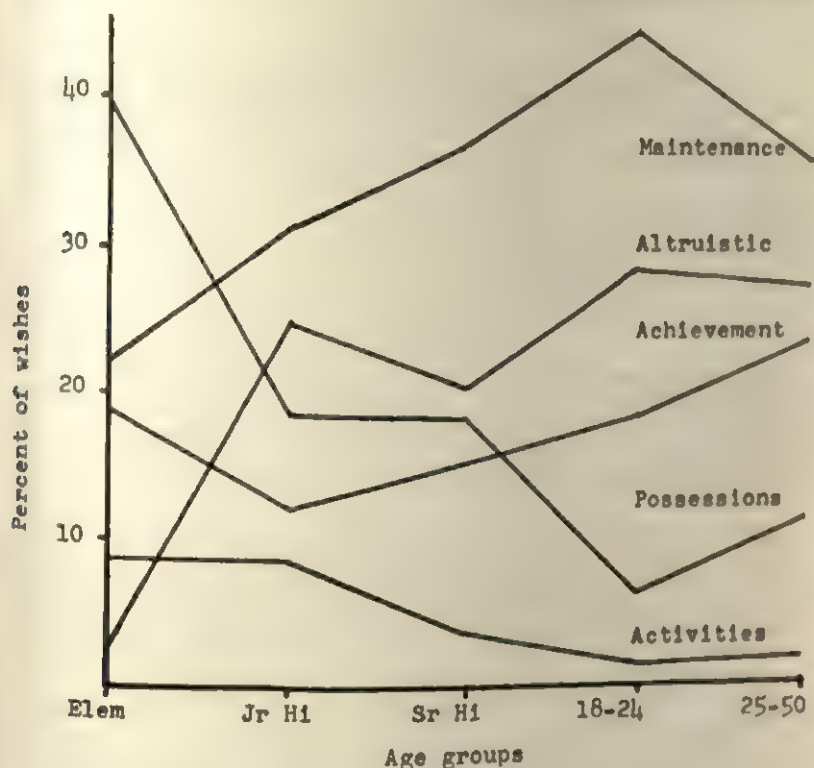


FIGURE 1  
PERCENT OF WISHES IN MAJOR AREAS BY AGE GROUPS

deliberately chose only enhancement, success, and recognition type wishes for inclusion in achievement categories.

Several categorical trends are of interest. Jersild's (7) comment on children's lack of wishes dealing with increased ability as a factor in achievement finds support in the data of Table 1, which further shows a gradual growth in the frequency of such wishes through middle adulthood. The gradual reversal in popularity of *love-have lover* and *marriage* from the youngest to the oldest group is of interest because of its predictability. The significant shift in altruism categories from concern with world peace and improving humanity in the 18-24 year old college group to high concern with the welfare of loved ones is felt to be of developmental interest, certainly supporting the hypothesis of developmental personality theorists who posit a period of "generativity" in motivational development at this age.

TABLE 1  
PERCENT OF WISHES BY AGE GROUPS FOR HIGHER FREQUENCY CATEGORIES

Category	Elem	Jr Hi	Age group		25-50	All
			Sr Hi	18-24		
<i>Possessions</i>	39	18	18	6	11	18
Wealth	10	8	11	6	11	11
Sports etc. equipment	9	4	1	0	0	3
Motor vehicle	5	3	4	0	0	1
Pets	9	2	0	0	0	1
Clothes	3	1	1	0	0	1
<i>Activities</i>	8	8	3	1	2	4
Sports-games-music	4	5	2	0	1	2
<i>Maintenance</i>	22	31	37	44	36	34
Happiness-satisfaction	2	3	6	10	10	6
Health-long life	1	2	3	8	9	5
Love-have lover	4	8	7	4	2	5
Marriage	1	2	7	11	6	5
Peer relations	3	2	2	2	1	2
Own family relations	2	4	1	1	0	2
Living conditions	3	3	3	1	0	2
<i>Achievement</i>	18	12	15	18	23	17
Vocation	6	6	3	6	7	5
Ability	4	2	4	6	8	5
Education	6	4	6	2	3	4
Success-recognition	2	0	2	4	4	2
<i>Altruistic</i>	4	24	20	28	27	21
World peace	2	12	10	10	5	8
Improve humanity	1	10	7	14	8	8
Welfare of loved one	1	2	3	5	14	5
<i>Miscellaneous</i>	9	7	7	2	1	6
More wishes	2	3	1	2	1	2
School avoidance	4	2	3	0	0	2
Other avoidance	2	1	2	0	0	1

Note:  $N = 300$  wishes for each age group and 1500 wishes for the "All" column.

TABLE 2  
PERCENT OF RESPONDENTS INDICATING HIGH CONFIDENCE AND A  
HIGH SENSE OF PERSONAL CONTROL

Attribute	Elem	Jr Hi	Age group		
			Sr Hi	18-24	25-50
High confidence	49	45	50	55	58
High sense of personal control	62	50	66	84	79

Note:  $N = 100$  for each group.

Table 2 shows the results from the exploration relative to age changes in confidence and the sense of personal control. None of the changes in confidence is significant, although a slow trend towards increased expectations of success with age may be observed to occur past the junior high group. Significant increases in the sense of personal control may be observed from junior high and from senior high to the 18-24 year group, suggesting that this construct and like constructs, such as internal locus of control, may indeed be significantly influenced by age. The twin increases of both constructs over this three age-group span may be observed and implications drawn relative to increased personality strength from early adolescence to early maturity. It is of interest to note the depressed percents for the junior high group in comparison to the other groups. It could be inferred that this age group does tend to lack personality characteristics often associated with personal competence and may reflect findings of other investigators relative to the high-stress conditions often faced in early adolescence.

Table 3 presents the variations in percents for major areas by sex. Junior high differences are most marked, indicating male preferences for possessions and activities and to a nonsignificant degree for achievement. The higher male interest in possessions trends into senior high, while the higher male interest

TABLE 3  
PERCENT OF WISHES BY SEX AND AGE GROUP IN MAJOR AREAS

Major area	Age group											
	Elem		Jr Hi		Sr Hi		18-24		25-50		All	
	M	F	M	F	M	F	M	F	M	F	M	F
Possessions	39	39	23	13*	21	15	6	6	12	10	20	17
Activities	9	7	11	4*	3	3	1	1	3	1	5	3
Maintenance	19	25	22	40*	35	39	43	46	37	35	31	37*
Achievement	22	15	15	10	15	15	23	13*	26	21	20	15*
Altruistic	4	5	19	30*	17	23	25	32	20	33	17	25*

Note:  $N = 150$  wishes for each sex in each age group and 750 wishes for each sex group in the "All" column.

\*  $p < .05$ .

for achievement exists in all but one of the age groups. Females may be seen to show a pervasively higher degree of concern for maintenance and altruistic areas, reaching significance levels in several age groups. The sex differences across all groups are of interest in that maintenance and altruistic wishes are clearly more frequent for females, with categories *marriage* and *welfare of loved one* showing the highest percents for females within the two areas (not shown). Achievement is of more general concern for males in general, with the significantly different categories being vocational achievement and increased ability. Within the exploratory data there were no significant sex differences within age groups; however, there was a significant trend indicated across three groups from junior high to the 18-24 year level for females to indicate a greater sense of personal control than males.

Table 4 presents the results from the elementary and junior high samples previously set aside as representing the lower socioeconomic condition. Comparison with the data used in establishing trends in the main study for middle class groups suggests that socioeconomic status is a significant variable influencing age trends. More frequent wishes for possessions occur at both elementary and junior high levels, the difference being significant in junior high; the socioeconomic differences is significant for boys at both levels. An achievement difference for elementary boys may be noted; lower socioeconomic boys are much less concerned with achievement than are their higher status counterparts. Altruism wishes decrease for the lower socioeconomic group in junior high school; this difference is significant for girls and for the group as a whole. Thus, in early adolescence desires for possessions seem higher and for altruistic events lower for the lower socioeconomic group. Although not shown in the table, it is of interest to report that lower socioeconomic boys, when compared to their high socioeconomic counterparts, showed a significantly lower frequency of high personal control scores in the elementary

TABLE 4  
PERCENT OF WISHES IN MAJOR AREAS BY SOCIOECONOMIC CONDITION AND SEX FOR ELEMENTARY AND JUNIOR HIGH AGE GROUPS

Major area	Age group and socioeconomic condition											
	Elem						Jr Hi					
	Middle			Lower			Middle			Lower		
	M	F	Total	M	F	Total	M	F	Total	M	F	Total
Possessions	39	39	39	54	34	44	23	13	18	36	21	28
Activities	9	7	8	5	11	8	11	4	8	6	5	6
Maintenance	19	25	22	17	29	23	22	40	31	23	35	29
Achievement	22	15	18	9	16	13	15	10	12	17	12	15
Altruistic	4	5	4	9	5	7	19	30	24	12	21	16



group, yet in junior high had a greater frequency approaching significance. Such a reversal was not suggested in the confidence measure. Lower socioeconomic girls showed higher frequencies in the confidence measure at both the elementary and secondary levels, but not to significance. It is reasonable to speculate that lower socioeconomic status tends to engender personality strength in junior high in comparison to middle class status.

#### D. DISCUSSION

One purpose of the present study has been to examine the effect of age on wishes and thus be able to make inferences about motives, interests, etc. It should be recognized, however, that there are several limitations to the accuracy of the method used. Cross-sectional sampling permits the possibility that significant generational values are retained from an earlier age, and that what seems to be an age difference alone is contaminated by effects from an earlier cultural milieu. It is also possible that such situational variables as the teacher, the atmosphere, the voluntary aspect of college attendance when compared to the public school groups, the time of the year, etc. are of greater importance than predicted. Finally, the subjective aspect of categorizing often ambiguous wishes must be considered as a factor in lessening the validity of the results. The findings relative to the exploratory data are subject to the same kinds of limitations and, in addition, are restricted in comparability to other findings because the operational definitions of "confidence" and "sense of personal control" used herein are unique.

In spite of the methodological limitations, it is felt that age has been a significant influence on these data, as witness their comparability with previously established trends and their conformity with developmental and sociological theory. It was demonstrated that wishes do tend to become more general, less materialistic, and more altruistic with age, although the age at which a principle becomes inoperative depends on the area or category taken as a measure. Wishes for specific possessions and activities do decrease up through early maturity; wishes for happiness and success increase. Wishes for world peace and improving humanity only show increases up to the junior high level, however. The categorization process undoubtedly masked the specific nature of some wishes, making a definitive statement impossible. The gradual growth in achievement wishes through middle maturity was shown as a previously unrecorded trend needing further substantiation. It is possible that the college student sample was heavily biased in this area or that the categorization process was too selective and that age as a variable is of little influence.

The data from the "confidence" measure did not show significant variation from chance expectations in any of the groupings used. It may be that the measure was too gross to show degrees of change, or it may be that confidence is indeed not an age-specific characteristic. Further, it is possible that the measure did not sufficiently tap actual expectations of success because of the relative unimportance of the wish rankings. On the other hand the changes in "sense of personal control" data were considerable, showing significant increases from junior high to the 18-24 year group, and indicating sex and socioeconomic variations. It is possible, therefore, that the sense of personal control was projected on to wishes and that the variations do reflect actual group deviations in this personality characteristic. Certainly the method and the findings are worthy of further investigation.

Sex and socioeconomic status were found significantly to influence various categories and areas in ways predictable from previous findings or from theory. The extension of the data into middle adulthood demonstrated that there may be limitations to previously established principles in age trends. Whether adulthood ages become stable in degrees of generality, materialism, and altruism, as somewhat indicated in the data, and whether age changes occur only in categorical sources of satisfaction, as suggested further, remain for future investigation to establish. However, it is indicated that caution should be used in extending wish data trends beyond adolescence.

It would have been of interest to have been able to extend the sampling to older age groups. It may be hypothesized that the adulthood plateau as discussed above does change at least in the later stages of adulthood, and that wishes could become more specific, more materialistic, and less altruistic, thus completing the developmental cycle.

A comparison of popular categories across studies is possible through reference to the Horrocks and Flory (6) study as reported in Horrocks (5). A comparison of the five most frequent categories of wishes from the middle class high school and college groups of the 1960 study with those of the present (1971) study is as follows:

1960: *High school*—wealth, happiness, health, success, world peace. *Early college*—health, marriage, success, world peace, serve humanity.

1971: *High school*—wealth, marriage, love-have lover, world peace, improve humanity. *Early college*—happiness, health, marriage, world peace, improve humanity.

It may be observed that the listings of the present study differ in two respects from the earlier study. The presence of *marriage* and *love-have lover* in the present study, and the absence of *success*, do suggest that the decade of

the 60's has produced some movement in motives for middle class adolescents away from achievement towards maintenance-oriented desires.

Continuing the low socioeconomic comparisons made by Horrocks and Flory (6), the following lists compare the five most frequent categories (unranked) of wishes in three low socioeconomic adolescent groups:

1932 [data from Washburne (8)]: wealth, car, clothes, happiness, school skill.

1960 [data from Horrocks and Flory (6) as cited in Horrocks (5)]: wealth, car, education, job now, world peace.

1971 (data from present study): wealth, car, vocation, world peace, improve humanity.

It should be observed that the adolescent groups are unequal in age span in that Washburne's sample covered students from 14 to 17 years, the Horrocks and Flory group were all juniors and seniors in high school (16-19 years), and the present study uses junior high students (12-16) as subjects. Despite this age variance, a considerable continuity may be seen across the three samples. Wishes in relation to a job and automobile are present in all three groups, whereas these wishes were not among the top five for the higher socioeconomic groups presented above. While the 1960 group added world peace to the 1932 selections, the 1971 sample added a second altruistic category and lessened concern for education over the 1960 group. It might be suggested that the decade of the 60's produced more concern among low status adolescents at least for altruistic categories and less for achievement. This suggested change in motives, together with the absence of success wishes for the middle class group, strongly suggests a generalized movement away from achievement orientations for adolescents.

### E. CONCLUSIONS

It has been shown that trends toward increased generalization, increased altruism, and decreased materialism continue to be valid principles for describing age changes in wishes, although the age at which change ceases varies with the measure and, in general, may not apply in adulthood. Age-specific changes in categorical sources of satisfaction were found to occur, however, and may represent the basis for studies of wishes in adulthood. The suggestion was made that future studies might find that late adulthood wishes again vary along the generality, altruistic, and materialistic dimensions, completing the developmental cycle. It was found that the measure of "sense of personal control" showed significant variation with age, sex, and socioeconomic condition, leading to speculation as to potential usefulness in future studies.

Sex and socioeconomic condition as variables, in addition to age, were found to be associated with significant wish changes. Comparison of results of previous studies was made, resulting in speculation that the decade of the 60's may have produced less achievement orientation among adolescents in the culture.

#### REFERENCES

1. ABLES, B. The three wishes of latency age children. *Devel. Psychol.*, 1972, **6**, 186.
2. BOYNTON, P. L. The wishes of elementary school children. *Peabody J. Educ.*, 1936, **13**, 167-174.
3. COBB, H. V. Role wishes and general wishes of children and adolescents. *Child Devel.*, 1954, **25**, 161-172.
4. GRAY, S. The wishes of Negro school children. *J. Genet. Psychol.*, 1944, **46**, 225-237.
5. HORROCKS, J. E. *The Psychology of Adolescence* (3rd ed.). Boston: Houghton Mifflin, 1969.
6. HORROCKS, W. B., & FLORY, E. Wishes of high school and college students. Unpublished study cited in Horrocks, J. E., *The Psychology of Adolescence* (3rd ed.). Boston: Houghton Mifflin, 1969.
7. JERSILD, A. T. Children's wishes, dreams, fears, daydreams, pleasant and unpleasant memories. *Psychol. Bull.*, 1933, **30**, 552-553.
8. WASHBURN, J. N. The impulses of adolescents as revealed by their written wishes. *J. Juv. Res.*, 1932, **16**, 193-212.
9. WHEELER, D. K. Expressed wishes of students. *J. Genet. Psychol.*, 1963, **102**, 75-81.

*Department of Psychology*  
*The Ohio State University*  
*Columbus, Ohio 43210*



## THE EFFECT OF FIRST NAMES ON CONFLICTED DECISIONS: AN EXPERIMENTAL STUDY\*<sup>1</sup>

*Temple University*

---

THOMAS V. BUSSE AND CRAIG LOVE

---

### SUMMARY

In a study of 80 fifth grade American boys and girls, an attempt was made to demonstrate a causal link between children's first names and judgments about the rightness or wrongness of those children's conflicted decisions. It was expected that children possessing liked names would be judged as right more frequently than children having disliked names. The expected effect was not found. Reasons for this negative finding were discussed.

---

### A. INTRODUCTION

Several studies (1, 6, 9) have shown that different first names carry different connotations. For example, Schoenfeld (9) found that the name Richard was thought of as "good looking." Other studies (2, 3, 4, 5, 7, 8) have found a correlation link between first name desirability and various personal and intellectual characteristics (e.g., personal popularity, emotional disturbance, IQ, reading achievement). Presumably, these findings suggest that first names are one of a series of "given" personal characteristics (e.g., eye color, hair color, ethnic background, physical stature) that influence individuals' perceptions of one another. However, no attempt has been made to show that a causal relationship exists between first name desirability and its supposed effects.

The present research attempted to demonstrate experimentally a causal link between children's first names and judgments about the rightness or wrongness of those children's decisions. Specifically, it was predicted that subjects would judge that characters possessing liked names in conflicted-decision stories were right in their choices significantly more than characters who had disliked first names.

---

\* Received in the Editorial Office on April 26, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The authors most appreciated Margaret Vogelsson's assistance in data analysis. George Byrne, Harry Finleson, and William Sanni are thanked for their help in obtaining subjects. Richard Mansfield and Margaret Vogelsson gave helpful comments on an earlier draft of this paper.



## B. METHOD

### 1. *Subjects*

Forty fifth grade boys and 40 fifth grade girls were randomly selected from a socioeconomically heterogeneous school. Six male and two female subjects were Negro; the remaining subjects were white.

### 2. *Stories*

Ten conflicted-decision stories were chosen from 22 initially composed by the authors. The final 10 stories were both chosen and revised through pretests with about 60 fifth graders at another school in the district. The selection of the stories was based on clarity and equality of right and wrong decisions. Approximate equality was necessary for the first names to have any reasonable possibility of affecting the choices.

The same 10 stories about elementary age children were individually read to boys and girls except that the sex of the characters was changed to match that of each subject. Each story concluded with the main character choosing between two courses of action. The subjects were then asked whether they thought the character was right in his (her) decision. A typical story was as follows:

"(Carol) was elected captain of the volleyball team. Thus, it was (Carol's) job to select the team members. (Carol) felt that she just had to win today's game, so she wanted to pick the best players. (Carol's) best friend wasn't a very good player but would be really hurt if she wasn't on (Carol's) team. However, (Carol) didn't choose her best friend. Was (Carol) right or wrong?"

### 3. *Techniques*

The first names of the story characters were randomly varied in the stories. Half of the first names were well-liked names, and half were disliked names. This changing of the first names of the story characters constituted the experimental manipulation of the study.

The liked and disliked names were obtained in a study done two years earlier with the then fifth graders in the same school district. This previous study involved 172 boys and 143 girls rating all of the first names of children in grades two through six in the same school district as the current study. Four randomizations of the boys' names and four randomizations of the girls' names were used in making this rating. In all, 179 different boys' names and 246 different girls' names were rated "like," "neutral," or "don't like."

Boys' names for the characters were used for the male subjects. Girls' names

were used for the female subjects. The 20 boys' names most liked by boys were used as the liked names for boys. The 20 boys' names most disliked by boys were used as the disliked names. The 20 most liked and 20 most disliked girls' names as rated by girls were likewise used.

The stories were presented in a single randomized order. Separately for boys and girls, the 20 liked and 20 disliked names were randomized over stories and subjects with the restrictions that each subject should have five disliked names in his 10 stories and that each of the 40 names should be used once in each of the 10 stories. In addition, no name was allowed to appear more than once in the stories of any subject. Nor could a subject's first name be used in the stories presented to him.

Two scores were calculated for each subject: (a) The number of times out of five that a subject judged as right a child with a liked name. (b) The number of times out of five that a subject judged as right a child with a disliked name.

### C. RESULTS

Nine of the 10 stories showed at least 25% of the subjects judging the character in the story to be right and at least 25% judging wrong. One story showed a 13% right, 87% wrong division. Since most of the stories fulfilled the pretest expectation of mixed judgments on the stories, a reasonable test could be made of the effect of the name manipulation.

The mean number of judgments of "right" for stories with liked names was 2.38 ( $SD = 1.02$ ) for boys and 2.50 ( $SD = 1.12$ ) for girls. The mean number of judgments of "right" for stories with disliked names was 2.30 ( $SD = 1.01$ ) for boys and 2.40 ( $SD = .97$ ) for girls. A judgment of "right" represents agreement with the action of the character in the story.

Wilcoxon's matched-pairs signed-ranks test was used to test the difference between means. Neither the test with boys ( $T = 236.5$ ,  $z = .23$ ,  $p > .10$ ) nor with girls ( $T = 181.5$ ,  $z = .49$ ,  $p > .10$ ) reached significance. Combining the data for boys and girls likewise did not yield significance ( $T = 822$ ,  $z = .48$ ,  $p > .10$ ). All were one-tailed tests.

### D. DISCUSSION

The expected name effect was not found. The decisions of children in the stories assigned liked names were not judged to be right significantly more than the decisions of children in the stories assigned disliked names. This first attempt to show a causal connection between a child's first name and other persons' perceptions of the child met with failure.

The present findings suggest that the correlations between first name desirability and, for example, personal popularity might not reflect a causal connection between them, but only a common tie to a third variable (or variables) that affects both of them. For example, perhaps mothers' personalities influence both the types of names to be chosen for babies and the popularity of the children as they grow older. A correlation would then exist between personal popularity and first name desirability. But first names would not therefore be a cause of the personal popularity.

It can also be argued that the present test of the hypothesis would have been supported only if first names have a broad, general effect on children's perceptions of one another. Perhaps, the causal effect of first names, if it does operate, is limited to personal popularity and similar variables.

#### REFERENCES

1. BUCHANAN, B. A., & BRUNING, J. L. Connotative meanings of first names and nicknames on three dimensions. *J. Soc. Psychol.*, 1971, **85**, 143-144.
2. BUSSE, T. V. First names, achievement, and IQ. Paper given at the Eastern Psychological Association meeting, Washington, D. C., 1973.
3. ELLIS, A., & BEECHLEY, R. M. Emotional disturbance in children with peculiar given names. *J. Genet. Psychol.*, 1954, **85**, 337-339.
4. HARTMAN, A. A., NICOLAY, R. C., & HURLEY, J. Unique personal names as a social adjustment factor. *J. Soc. Psychol.*, 1968, **75**, 107-110.
5. HOUSTON, T. J., & SUMNER, F. C. Measurement of neurotic tendency in women with uncommon given names. *J. Gen. Psychol.*, 1948, **39**, 289-292.
6. LAWSON, E. D. Semantic differential analysis of men's first names. *J. of Psychol.*, 1971, **78**, 229-240.
7. McDAVID, J. W., & HARARI, H. Stereotyping of names and popularity in grade-school children. *Child Devel.*, 1966, **37**, 453-459.
8. SAVAGE, B. M., & WELLS, F. L. A note on singularity in given names. *J. Soc. Psychol.*, 1948, **27**, 271-272.
9. SCHOENFELD, W. N. An experimental study of some problems relating to stereotypes. *Arch. Psychol.*, 1942, **38**(No. 270), 1-57.

*Department of Educational Psychology*  
*Temple University*  
*Philadelphia, Pennsylvania 19122*

## LANGUAGE DOMINANCE AND BILINGUAL RECALL\*

*Indiana University*

---

STARRETTE DALTON

---

### SUMMARY

Both unilingual and bilingual word lists were recalled by 98 bilingual subjects. The direct relationship between dominance and recall which occurs under unilingual conditions was found to disappear under bilingual conditions. Bilingual lists (*a*) were recalled as well as unilingual lists and (*b*) led to superior recall of Spanish for nearly all subjects.

---

### A. INTRODUCTION

Tulving and Colotla (8) reported that impairment in recall of bilingual and trilingual lists is largely accounted for by impairment in recall of the language recalled best when presented alone. This is tantamount to a reversal of dominance between unilingual and bilingual recall conditions; i.e., the language recalled best from a unilingual list is recalled second best from a bilingual list. This result is so counterintuitive that further examination of its occurrences was undertaken. In addition, an attempt was made to determine whether unilingual or bilingual lists are more difficult to recall. Previous investigators (5, 6) have found that bilingual lists with no obvious semantic categories are recalled as well as unilingual lists. Tulving and Colotla found multilingual lists more difficult to recall and maintain that the issue is still unresolved.

### B. METHOD

#### 1. *Subjects*

All 98 Ss were literate in both English and Spanish and were students at Pan American University. The 67 females and 31 males were volunteers tested in the classroom after all students unable to read and write both Spanish and English were excused.

#### 2. *Materials*

Words were selected from the 324 high frequency English nouns which Tulving and Colotla used in their study of trilingual free recall. With other

---

\* Received in the Editorial Office on April 27, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.



factors held constant, a list of high frequency words is better recalled than one of low frequency (1). Tulving and Colotla seem to have committed an oversight when they attributed their subjects' superior recall of English to language imbalance. Many of the Spanish translations were so infrequent as not to appear in a frequency dictionary of Spanish words (3). With this in mind, only words that (a) occur more than 35 times per million in both English and Spanish and (b) clearly belong to only one of the two languages were selected. After this selection procedure 119 words remained in the pool. To estimate the difficulty level of each word in the free recall situation they were randomly assigned to one of three 39-item lists and administered in free recall. The English translations were administered to 22 English monolinguals, and the Spanish translations of the same three lists were administered to 16 Spanish monolinguals. Then four word lists, each 16 words long, were formed by randomly assigning words to one of four lists and making the appropriate translation if necessary to yield one Spanish (S), one English (E), and two mixed (M[1] and M[2]) lists. None of the first four or last four words in the 39-word lists were used, since recency and primacy could elevate their  $p$  levels beyond levels that would be expected with recency removed and order of appearance altered. Average  $p$  level of each list was computed, and substitutions were made in the Spanish and the first mixed list (M[1]) to bring their average  $p$  levels down to the level of the other two lists. Within each mixed list the average  $p$  level of the eight English and the eight Spanish words was also equated. The resultant  $p$  levels for English, Spanish, M[1], and M[2] were .2864, .2896, .2790, and .2880, respectively. They were, of course, different when computed after administration of the 16-word lists: (a) The lists were shorter, and this raised the  $p$  levels. (b) The presentation rate was faster (to avoid ceiling effects), and this decreased the  $p$  level.

### 3. Procedure

Subjects were tested in small groups (four to 10 Ss each), randomly assigned to one of the four experimental groups (20 Ss each) and presented with four 16-word lists reflecting three language conditions: Spanish (S), English (E), or mixed (M). A four by four Latin Square was used for randomization purposes. Actually 98 Ss were tested and used in most analyses. When equal cell size was necessary for computational purposes Ss were removed randomly to provide an  $n$  of 20 in each experimental unit. Each cell order of presentation was randomized separately.

A pre-experimental questionnaire was given to each subject concerning where and how he learned each language, his relative strength in each lan-



guage, his classification as a coordinate or compound bilingual, and the language skills of his parents. The word lists were presented to the Ss on 35 mm slides by means of a Kodak Carousel projector and an electronic timer to regulate the rate of presentation. One word was presented each 4 seconds with an approximate exposure time of 3.5 seconds. After each list of 16 words, a slide with a complex geometric figure was presented for 20 seconds (Ss were required to copy the figure until it disappeared) to remove any variance introduced by the recency effect (7). To avoid a set in either language, instructions were given to Ss in both English and Spanish; one half received English first and the other received Spanish first. In short, the instructions were to observe each list, to trace the geometric figure until it disappeared, and to write, in any order, all the words they could remember. Amount of time given for recall was 128 seconds (twice the presentation time). Two minute rest periods were allowed after each list to reduce interference among lists.

A double classification, repeated-measure analysis of variance was carried out to compare recall of the different lists and of the different groups. Analyses for simple effects and multiple comparisons were also carried out to determine exactly where the differences were located.

### C. RESULTS AND DISCUSSION

#### 1. *Efficiency of Bilingual Free Recall*

The four lists differed in difficulty,  $F(3,228) = 14.4$ ,  $p < .01$ , with means of 8.42, 7.19, 8.13, and 7.08 for E, S, M[1], and M[2], respectively. The four orders of presentation (or groups) were significantly different,  $F(3, 76) = 5.7$ ,  $p < .01$ , as was the Group  $\times$  Lists interaction,  $F(9, 228) = 3.64$ ,  $p < .05$ .

Since the particular concern here is whether the mixed lists are more difficult than the unilingual ones, the differences among the lists were examined with the Newman-Keuls procedure. The mixed lists did not differ in difficulty from the unilingual ones in any systematic fashion. The means of the combined unilingual and bilingual lists were 7.8 and 7.6, respectively. The mixed lists differed from each other as did the unilingual lists ( $p < .05$ ). Among those factors contributing to list difficulty, the bilingual nature of the list did not seem to figure prominently. At least one other factor played an important role because M[1] and M[2] differed significantly. One can explain the apparent dominance of English under unilingual conditions as due to the sample of bilinguals selected.

Since the Group  $\times$  List interaction was significant, the variation due to the simple main effects of each factor was examined under each separate condition

of the other factor. For three of the four groups tested the four lists were unequal in difficulty ( $p < .01$ ). The fourth group, as depicted in Figure 1, demonstrates the same relative scoring pattern as the other groups although the differences were insignificant. This may be due to the fact that the fourth group had the easier lists presented last. This could have counteracted the proactive inhibition,  $F(3, 237) = 6.72, p < .01$ . This reasoning is supported by group two: the fact that the two most difficult lists are in last position combined with the effect of proactive inhibition produced the greatest list variance and hence the highest  $F$  value,  $F(3, 228) = 14.3, p < .01$ . Each list differed in difficulty across the four groups ( $p < .05$ ). Apparently no large difference due to one list can account for the differences among groups.

In terms of Tulving's theory of memory this study suggests that the storage of language markers does not occupy space in memory which would reduce the number of words that can be recalled. Other factors must be much more important than space in memory occupied by language markers or organization provided by these markers because the difference between the two mixed lists is much greater than the difference between the unilingual and the bilingual lists. This indicates that further study of the specific character-

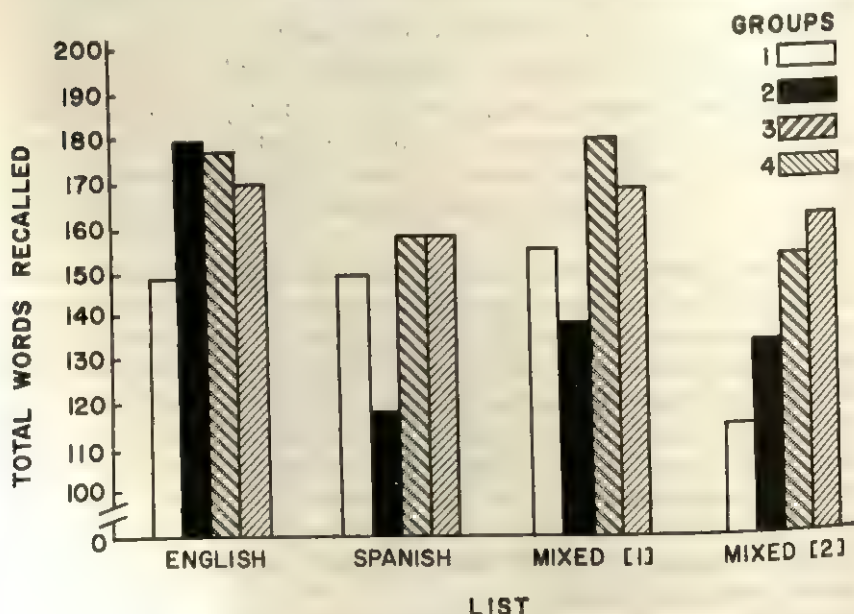


FIGURE 1  
PERFORMANCE OF SEPARATE GROUPS ON DIFFERENT LISTS

istics of the lists and of the sample used in warranted. Perhaps once these extraneous factors have been isolated, we can talk more meaningfully about the analogy of space and memory.

## 2. Dominance

Two ways of defining language dominance will be considered: *evinced dominance*—a subject is classified as dominant in the language he recalls best under unilingual conditions; and *rated dominance*—a subject is judged dominant in the language he rates as his strongest.

*a. Evinced dominance.* Tulving and Colotla reported a reversal of evinced dominance under unilingual and bilingual conditions of recall, and the present experiment, at first, was thought to replicate this counterintuitive discovery. But since 24 Ss recalled Spanish best under unilingual conditions, the impairment in their bilingual recall should have resulted mainly from an impairment in Spanish. Examination of the data, however, reveals that nearly all subjects were greatly impaired in their recall of English under bilingual conditions (Figure 2). Only three subjects recalled English better than

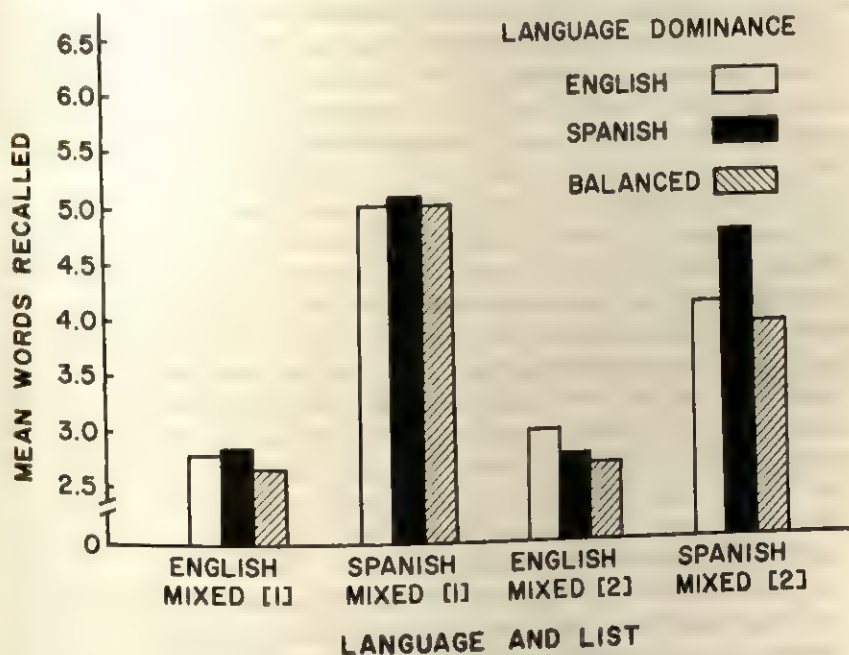


FIGURE 2  
EVINCED DOMINANCE AND PERFORMANCE UNDER BILINGUAL CONDITIONS

TABLE 1  
LANGUAGE DOMINANCE AND PRESENTATION CONDITIONS

Ss' dominant language	Condition	
	Bilingual	Unilingual
Spanish	79	24
English	3	59
Balanced	16	15
Total	98	98

Spanish under bilingual conditions (Table 1), and all three subjects were dominant in English under unilingual conditions. Thus all of those who had evinced dominance in Spanish actually had less impairment in their stronger language. Regardless of dominance under unilingual conditions, 95/98 Ss recalled Spanish better under bilingual conditions.

The phenomenon yet to be explained is the suppression of English and enhancement of Spanish recall under bilingual conditions. Possibly the Ss had a set for English, since they were tested in classrooms where most instruction was provided in English. This would highlight the distinctiveness of Spanish words, and results could be analogous to the Von Restorff effect (2); i.e., Spanish words in an English setting would be unique items in an otherwise homogeneous list. However, for a sequence of words such as EEES one would predict greater recall of the first Spanish word, since it should be more distinctive than the second. The present data do not support this prediction.

Kolers (4) maintained that switching from one language to another, at least when reading a mixed passage, occupied time. It was thought that switching from Spanish to English could be more costly (in time) than switching from English to Spanish. Lowered recall of English in the bilingual condition could be interpreted as a result of the greater difficulty of switching from Spanish to English. More time would be spent switching and less time encoding the English words which would lower the English recall. But closer examination of the data for each group contradicts this hypothesis. In a sequence such as SSSEE one would expect the switch to be complete on the first English word, and the recall of the second English word not to be impaired by a switch. This tendency was not found in the data; i.e., the second English word was equally impaired.

*b. Rated dominance.* Regardless of evinced dominance Ss tended to rate themselves as dominant in English (Table 2,  $\chi^2 = 8.23$ ,  $df = 2$ , ns). The inclination to rate themselves as dominant in English may reflect the fact that English is the socially most prestigious language in the community, and Ss



TABLE 2  
SELF-RATED AND EVINCED DOMINANCE

Dominance measure	Spanish	Number of subjects		Total
		English	Balanced	
Self-rated	9	70	19	98
Evinced	24	59	15	98

prefer to view themselves as stronger in English than their evinced dominance scores indicate. This may account for the discrepancy between dominance in the unilingual and the bilingual conditions. Under unilingual conditions subjects could more readily distinguish Spanish recall from English recall and hence concentrate on recalling the latter, the socially preferred language. Under bilingual conditions separation of English from Spanish recall would be more difficult, and the unbiased strength of each language might more clearly be revealed. Therefore, English would appear as the dominant language of most Ss under unilingual conditions where prevailing social attitudes toward the different languages can more easily influence recall behavior. It is hypothesized that Spanish would be recalled best by nearly all Ss under bilingual conditions in which the languages are thoroughly mixed, so that the effort to separate them in recall is too great and subjects are reduced to recalling any word they can despite the language in which it appears. A test of the present hypothesis would be relatively simple in a community where Spanish is most prestigious. If, under these circumstances, Spanish is not recalled better than the less prestigious language, other interpretations of the emergence of Spanish phenomenon found in the present study could be set fourth and tested.

In light of the two measurements of dominance, a final piece of evidence is available which more clearly contradicts the reversal of dominance hypothesis because the confusion concerning the different and somewhat contradictory measures of dominance is absent. Five Ss rated themselves as dominant in Spanish and also evinced dominance in Spanish under unilingual conditions (Table 2). All of these Ss recalled Spanish better under bilingual conditions. This more conclusively indicates that there is no reversal of dominance between unilingual and bilingual conditions.

### 3. Spelling Errors

There were more spelling errors in the Spanish list than in the English list and more spelling errors in the Spanish halves of the mixed lists. Obviously spelling errors were more likely in Spanish regardless of list difficulty or condition of presentation. This is consistent with the subjects' preponderance of practice with written English. Of greater interest, however, is the relationship



between spelling errors and difficulty of recall from the different lists (Figure 3). The four lists differed in number of spelling errors,  $F(3, 237) = 3.9, p < .01$ . The two most difficult lists had more spelling errors than the two easiest lists ( $p < .01$ ).

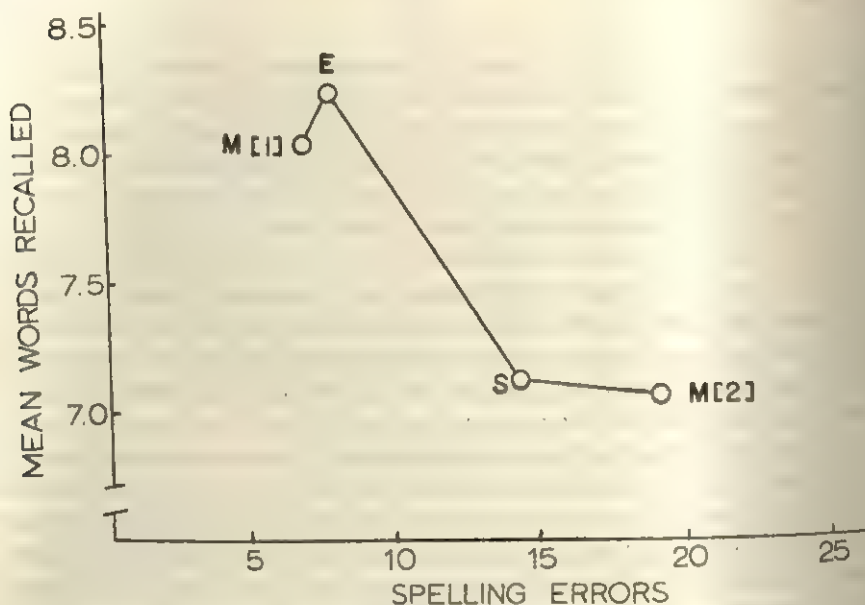


FIGURE 3

RELATIONSHIP OF MEAN WORDS RECALLED AND SPELLING ERRORS

Definitions of symbols for word lists are as follows: E = English, S = Spanish, M[1] and M[2] = mixed.

Tulving and Colotla have maintained that differences in recall are due, not to storage, but to difficulties of retrieval. The direct relationship between spelling errors and difficulty of recall does not deny the importance of retrieval, but indicates that the nature of the reproduction process which follows retrieval is also very important.

The relationship between spelling difficulty and amount recalled from each list may indicate that decoding of stimuli plays a more important role in determining amount recalled than has been previously recognized. Perhaps less emphasis should be placed on stimulus presentation and conditions of memory and more on the response conditions. This points to many interesting research possibilities. For example, would relative differences remain constant if recall were oral rather than written? In written recall, difficulties in decod-

ing an item may be magnified when Ss are given a limited amount of time for recall and must devote great effort to recalling as many items as possible. They don't have time to focus on correct spelling by simply checking their work or applying spelling rules which are habitually applied when writing. In oral recall this would be unnecessary.

Subjects were classified as coordinate or compound by means of information obtained for each subject concerning where and how each language was learned. In the case where the evidence was equivocal the subjects' self-classification as compound or coordinate was also taken into consideration. Several analyses were carried out to determine if the groups differed systematically. It was thought that a person who learned the two languages in distinct and separate contexts might tend to maintain this separation and divide a bilingual list into two sections according to language. This was not the case. In fact, no major differences could be found between the two groups with regard to the free recall task. Perhaps the free recall situation is insensitive to differences between compound and coordinate bilingualism.

At present, we cannot say definitively that bilingual lists are recalled as well as unilingual ones, but the evidence does seem to point in this direction. We can, it appears, dismiss the reversal of dominance phenomenon and concentrate on the more interesting emergence of Spanish under bilingual conditions.

#### REFERENCES

1. DEESE, J. Frequency of usage and number of words in free recall: The role association. *Psychol. Rep.*, 1960, **77**, 337-44.
2. DEESE, J., & HULSE, S. H. *The Psychology of Learning*. New York: McGraw-Hill 1967.
3. JUILLAND, A. G., & CHANG-RODRIGUEZ, E. *Frequency Dictionary of Spanish Words*. The Hague, The Netherlands: Mouton, 1964.
4. KOLERS, P. A. Bilingualism and information processing. *Sci. Amer.*, 1968, **218**(3), 78-86.
5. LAMBERT, W. E., IGNATOW, M., & KRAUTHAMER, M. Bilingual organization in free recall. *J. Verb. Learn. & Verb. Behav.*, 1968, **7**, 207-214.
6. NOTT, C. R., & LAMBERT, W. E. Free recall of bilinguals. *J. Verb. Learn. & Verb. Behav.*, 1968, **7**, 165-171.
7. POSTMAS, L., & PHILLIPS, L. W. Short-term temporal changes in free recall. *Quart. J. Exper. Psychol.*, 1965, **17**, 132-138.
8. TULVING, E., & COLOTILA, V. A. Free recall of trilingual lists. *Cognit. Psychol.*, 1970, **1**(1), 86-98.

Bureau of Educational Studies and Testing  
 School of Education  
 Indiana University  
 Bloomington, Indiana 47401



## SCALING MEANINGFULNESS (M) OF TRIGRAMS WITH CHILDREN AND RETARDATES\*<sup>1</sup>

CECIL O. CAMPBELL AND ALFRED A. BAUMEISTER<sup>2</sup>

---

*Department of Psychology, University of Alabama*

---

### SUMMARY

The production method and paired-comparisons procedures were used to scale 25 trigrams for meaningfulness with children and retardates. Reliabilities of scale values derived from subgroup comparisons were moderate. Test-retest reliabilities were somewhat higher. Correlations of these scale values with values obtained from adults ranged from .60 to .75. Adult meaningfulness values predicted paired-associate learning success for both children and retardates better than did production and paired-comparison values. The results support the assumption that meaningfulness norms obtained from adults are, under some conditions, appropriate for use in learning experiments with children and retardates.

---

### A. INTRODUCTION

During the past several decades, many investigators have reported meaningfulness (M) values for CVC trigrams scaled with adult subjects (1, 5, 8, 9, 12, 13, 14, 15). Normative data as extensive as these have not been collected for young children and mental retardates. When studying the role of meaningfulness in verbal learning by these subjects, investigators have typically selected verbal materials on the basis of adult norms (e.g., 2, 11, 17). The obvious assumption here is that meaningfulness values of words generalize, at least in an ordinal sense, from adults to children.

Palermo, Flamer, and Jenkins (16) made a somewhat tangential test of the validity of the assumption that adult norms may be used with children. They evaluated M values for adults and fifth grade children by inference from performance on a common paired-associate task, concluding that adult M values have validity for use with children. A more definitive test, however,

---

\* Received in the Editorial Office on April 30, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Supported by PHS research grant HD 02588.

<sup>2</sup> Direct requests for reprints to the second author at the address shown at the end of this article.

requires that normative data for children be collected and compared with norms obtained from adults. Only two investigators have provided data in this respect. One such study was conducted by Shapiro (18) who obtained association values for 52 CVCs from Noble's (14) list. Subjects were 600 children, 300 boys and 300 girls equally distributed in grades 4, 6, and 8. Noble's (13) production method was employed, and subjects were given 18 seconds to write associations. The high and low meaningfulness values computed from the children's responses were found to be significantly different when the values were dichotomized according to Noble's *m'* level. Rank-order correlations with values reported by Noble (14) ranged from .52 to .66. As Gaeth and Allen (4) have noted, such relatively low correlations tend to cast doubt upon the assumption that association values for children and adults are highly similar.

However, Shapiro's results may not adequately represent the range of scaled meaningfulness of CVCs for children because of artificial ceiling effects. Subjects were limited to a maximum of five different associations for each trigram during the relatively short period of 18 seconds. If the number of associations had been allowed to vary freely, a wider range of values may have obtained, and correlations with values obtained from adults may have been influenced. Also, Shapiro scaled only CVCs that were real words, and consequently, the entire range of Noble's values was not represented.

More substantial correlations between M values for children and for adults were obtained by Gaeth and Allen (4). Employing Archer's (1) procedures, these investigators obtained association values for several lists of trigrams for children in grades 4-6. Correlations between scaled M for children and adults ranged from .83 to .87. These results support the assumption that children and adults tend to rank-order words in the same way.

There has been no systematic attempt to scale meaningfulness of verbal materials for mentally retarded subjects. Therefore, in addition to providing additional data to test the contention that adult meaningfulness values are similar to those for children, the purpose of the present study was to provide similar data for retarded subjects. Another purpose was a comparative evaluation of the two scaling methods: namely, the production method (13) and paired-comparisons. The latter technique, which has not been previously employed to scale meaningfulness, was selected as an alternative to traditional scaling methods which might prove to difficult for immature subjects (e.g., responding to a five-point rating scale). There is some evidence, from verbal mediation research, that young children and, particularly, mental retardates experience a production deficiency (3). The production method of scaling M,



therefore, might yield different results compared to some method that does not require the subject to generate overt verbal responses.

## B. PHASE I

This study was conducted in two phases. In the first, normative data were collected for children and retardates. A major interest was the correlation between *M* values reported by Noble (14) and those obtained here for children and for retardates. Another focus was upon the development of scale values of meaningfulness by the paired-comparisons method.

### 1. Method

*a. Subjects.* Sixty-one mental reardates and 50 normal children served as subjects. Ten retarded and two normal subjects were excluded because of inability to recognize the letters of the alphabet. The data for three additional retardates were discarded as a result of failure to follow paired-comparisons instructions. The remaining subjects were 48 retardates and 48 normal children who were, within each intelligence group, randomly assigned to one of two groups receiving instructions for either the production method or paired-comparisons.

The retardates were selected from the population at a state residential institution. The names of residents who met appropriate criteria (*IQ* 50-80, *CA* < 30) were arranged in a random order according to a table of random numbers. The normal children were from grades 2-6 of a local elementary school. Selection was random with the restriction that mean *CAs* and *SDs* closely approximated mean *MA*s and *SD*s of the respective retardate groups. Mean *MA*s for the two retardate groups were 8.8 and 9.1 years. Mean *CAs* for the two normal groups were 8.7 and 9.1 years. Corresponding *SD*s were all approximately 1.1 years. Mean *IQ*s of the retardate groups were 60.0 and 58.9 with *SD*s of 7.7 and 6.6 for those receiving production method and paired-comparisons instructions, respectively.

*b. Materials.* Twenty-five CVC trigrams were selected on the basis of Noble's norms (14). Included were 20 nonwords and five words. Meaningfulness (*m'*) values ranged from 1.00 to 4.22, representing approximately the full range of Noble's values. The trigrams were selected at random with the restriction that they were equal scale units apart. Selection was further restricted in that only those trigrams that would be relatively easy to pronounce were included.

For the production method, trigrams were printed in 1¼ inch black letters on a 3 × 5 inch white cards. Guilford's (7) shortcut method for paired-

comparisons was used with five of the trigrams (FEG, JIT, MUS, CON, and BOY) serving as standards, thus separating the list into approximately equal intervals. Each standard was paired with every other trigram in the list for a total of 110 pairs which were printed on  $5 \times 8$  inch white cards with three inches of blank space separating the two members of a pair.

*c. Procedure.* The subjects in the production method groups were assigned to a scaling procedure that is essentially the technique devised by Noble (13), with the exception that the experimenter recorded the associations elicited by a trigram within a time limit of 60 seconds. The task should be less difficult for immature subjects if they do not have to write responses themselves. The procedure also allowed subjects to give a greater number of associations, since they did not have to take the time required to write them down. Subjects also did not have previously written responses in front of them to elicit chain associations. The subjects in the paired-comparison groups were presented the series of comparisons and were given a maximum of 10 seconds in which to point to the more meaningful trigram. Again, responses were recorded by the experimenter.

The 25 cards used with the production method were reshuffled for each subject in order to minimize any contextual effects. Two orders of the paired-comparison list were employed. The orders were essentially random with the exception that no trigram or standard was presented on any two consecutive pairs. This procedure was followed in order to minimize the possibility of selecting an item on the basis of having just selected or rejected it. To control for possible position effects, a table of random numbers was used to assign standards to left or right position within a pair.

Subjects were taken individually into a small room and seated across from the experimenter. Each subject was given a standard set of instructions appropriate to the specific scaling technique. In order to obtain reliability measures for the two scaling procedures, the 24 subjects in each group were randomly divided into two groups containing 12 subjects each. The scaling procedures were replicated within two months of the initial scaling to obtain an additional test-retest reliability measure for the two retardate groups.

## 2. Results

Scale values for the production method were obtained by calculation of the mean number of associations to each stimulus. Paired-comparison scale values are normal deviate transformations of the mean proportions of frequencies that each stimulus was selected as being more meaningful than comparison stimuli. The resulting scale values for both methods are presented in Table 1. The trigrams are arranged in the order that they appeared in Noble's (14)

TABLE 1

MEANINGFULNESS VALUES OF 25 SELECTED CVC TRIGRAMS SCALED FOR CHILDREN  
AND RETARDATES BY PRODUCTION METHOD AND PAIRED-COMPARISONS

Trigrams	Production method		Paired-comparisons	
	Normals	Retardates	Normals	Retardates
CIJ	1.88	.92	.00	.00
GOJ	2.17	1.58	.72	.64
GAX	2.33	1.58	.90	.59
ZUB	2.13	1.21	.52	.64
FEX	2.25	1.25	.85	.39
FEG	2.33	1.38	.46	.64
KIZ	2.13	1.33	.41	.54
YAW	2.29	1.29	.80	.41
TUK	2.79	2.08	.74	.69
CUZ	2.04	1.42	.59	.39
HEK	1.79	1.33	.90	.46
DAK	2.83	2.17	.64	.89
JIT	2.75	1.00	.64	.49
SUZ	2.25	1.29	.57	.56
NIC	3.08	1.71	.90	.59
MUS	2.63	1.29	.80	.59
POM	2.88	1.96	1.08	1.19
NAP	2.63	2.21	1.08	.92
RAW	2.67	1.75	1.00	.67
BIL	2.96	1.83	1.16	.61
HOT	4.54	3.21	1.44	.84
CON	2.83	2.00	.85	.59
GUN	3.92	2.92	1.44	1.00
CAL	3.08	1.83	.92	.84
BOY	3.71	2.46	1.10	.92

list. The range of values for the production method is 1.79 to 4.54 for children and .92 to 3.21 for retardates. Corresponding *SDs* are .31 and .24.

Pearson *r*s were calculated across subgroups as measures of scale value reliabilities. With the normal children, reliabilities of the production method and paired-comparisons were .70 and .64, respectively. Corresponding coefficients computed on the retardate data were .51 and .73 ( $p < .01$ ). Although paired-comparisons, relative to the production method, produced higher scale value reliability within the retardate sample, this difference did not reach statistical significance ( $p < .10$ ).

Test-retest reliabilities were also obtained for the retardates. Correlations were .77 and .75 for the production method and paired-comparisons, respectively. The two methods appear to produce scale values that are fairly stable over time.

Correlation coefficients relating the various scale distributions of *M* are presented in Table 2. Correlations of Noble's *m'* with scale values obtained for children were more substantial than those reported by Shapiro (18), but less than those observed by Gaeth and Allen (4). For the retardates the co-

TABLE 2  
PRODUCT-MOMENT COEFFICIENTS FOR COMPARISONS OF PRODUCTION METHOD AND  
PAIRED-COMPARISONS: NORMAL, RETARDATE, AND NOBLE'S  $m'$

Comparison	$r$
Normal Production vs. paired-comparison	.76**
Retardate production vs. paired-comparison	.74**
Normal vs. retardate production	.88**
Normal vs. retardate paired-comparison	.68**
Noble $m'$ vs. normal production	.75**
Noble $m'$ vs. normal paired-comparison	.73**
Noble $m'$ vs. retardate production	.67**
Noble $m'$ vs. retardate paired-comparison	.60**

\*\*  $p < .01$ .

efficients were slightly slower. Even in the case of the children, adult scale values of  $m'$  account for only half of the variance in meaningfulness of the trigrams.

Analysis of variance was performed on the production data by means of a Lindquist Type I design (10) with trigrams as repeated measures. A significant effect for Intelligence Group,  $F(1, 46) = 7.32$ ,  $p < .01$ , indicated that normal and retarded subjects differed with respect to the values assigned to the trigrams. Production scores were consistently higher for the children than for the retardates. The two distributions are highly correlated ( $r = .88$ ) as indicated in Table 2. The relative meaningfulness of the trigrams, then, was essentially the same for the two intelligence groups. Moreover, within each intelligence group the production and paired-comparison values are correlated substantially (.76 and .74). The paired-comparison values for the two intelligence groups do not correlate as highly as production values, but the correlations are still fairly high. Correlations with Noble's values are also substantial, ranging from .60 to .75.

### 3. Discussion

Both the production and paired-comparisons procedures yield fairly reliable scale values of meaningfulness for retardates and children. Moreover, these values are rank-ordered in much the same way for children and retardates and adults, as well. However, the fact should be stressed that, relative to children, retardates exhibited a production deficiency. For all trigrams, words and nonwords alike, the normal children gave more responses. The same conclusion, although not directly tested here, probably holds true for children relative to adults. The research and interpretative implications of these production differences are concealed in the correlational analyses that have been done previously with normative data.



One implication of these results is that normative *M* data are interchangeable between groups depending upon the purpose to which the data are to be put. If the application requires only that the words or trigrams hold their relative rank (as in Phase II of this study), then one may, with some justification, interchange normative values between different subject populations. But if an inference is to be drawn that depends upon absolute *M* value, such as in certain mediational contexts, then it probably is less appropriate to use adult *M* values with either children or retardates. In fact, the results reported here indicate that retardates exhibit a production deficiency relative to mental-age matched normals.

Although paired-comparison procedures have not been used previously to scale verbal materials, the technique may be a promising one with immature or handicapped subjects. The paired-comparisons method also appears to possess some technical advantages over the production method for scaling verbal materials. These results, although limited to trigrams, indicate that paired-comparisons produce fairly reliable scale values of *M* that are stable over time.

## C. PHASE II

### 1. *Introduction*

The purpose of the second phase was to determine the validity of the ordering of the scale values obtained previously. This was accomplished by evaluation of the success of the various scale values, including those for adults, in predicting paired-associate learning performance.

### 2. *Method*

*a. Subjects.* Twenty-four retardates and 24 normal children were selected in the same manner and from the same sources as described previously. Data for an additional 11 retardates were discarded because of failure to learn the task. The 24 subjects within each intelligence group were randomly assigned to one of two paired-associate tasks for a total of four groups consisting of 12 subjects each. The means of *MA* for the retardates were 9.3 and 9.4 years, and *CA* for the normals were 9.6 and 9.8 years. Corresponding *SDs* were 1.3, .98, .97, and 1.4. Mean *IQs* for the two retardate groups were 63 and 63.8 with *SDs* of 6.4 and 8.8.

*b. Materials.* The response terms in each of four six-item paired-associate lists were trigrams selected on the basis of meaningfulness values collected in the first phase described previously. The scale values in each distribution were transformed into normal curve deviates, and, on the basis of these scores, an



attempt was made to construct two paired-associate lists for each intelligence group that were comparable in meaningfulness value. (In construction of the lists one trigram was included in both retardate lists, and two in both normal lists.) The trigrams in each list were randomly paired with single-digits which served as stimulus terms. The four lists are presented in Table 3 along with learning scores. A Kodak Carousel projector was used to present the verbal materials which were typed on  $2 \times 2$  inch slides in capital letters. Rate of presentation was controlled by an interval timer.

*c. Design.* There were four separate paired-associate tasks, two for each intelligence group. Pilot data indicated that no more than six items could be learned by the retarded subjects within a period of 30 minutes. Two tasks were used so that learning scores could be obtained for more than six trigrams. In order to prevent serial learning, four random orders of each lists were employed. Subjects learned the paired-associate items by the anticipation method.

*d. Procedure.* Subjects were taken individually into an experimental room and seated in front of a screen. On the familiarization trial, subjects were asked to pronounce the individual letters of the trigrams and name the digits. Subjects were then given standard instruction for the anticipation method of paired-associate learning. The presentation rate of 5:5 seconds was automatically controlled. There was a 30 second intertrial interval during which the investigator told the subject that he was doing well. Criterion of learning was one errorless trial.

TABLE 3  
PAIRED-ASSOCIATE LISTS AND LEARNING DATA

List	Retardates		Normals	
	Paired-associate items	Mean errors to criterion	Paired-associate items	Mean errors to criterion
1	7-DAK	6.33	2-BOY	1.42
	3-POM	6.75	5-NIC	4.83
	1-GOJ	7.92	3-POM	3.00
	6-BOY	1.58	8-FEG	5.50
	9-ZUB	7.75	1-GOJ	5.25
	2-JIT	5.83	4-HEK	5.25
2	7-DAK	7.75	5-GUN	2.58
	1-HOT	1.08	3-POM	3.08
	4-TUK	9.50	1-CAL	3.42
	9-MUS	4.92	9-MUS	3.83
	6-YAW	9.92	2-SUZ	4.42
	5-CUZ	7.50	8-FEG	4.83

### 3. Results

The results of *t* tests performed on the learning data revealed that, for each intelligence group, the two lists were comparable in learning difficulty. For purposes of analyses, the lists were combined. The learning score for each trigram was mean errors to criterion. Each of the sets of learning scores was correlated with the production and paired-comparison values for the trigrams scaled for each intelligence group and with values for adults obtained by Noble (14). The correlations are presented in Table 4. Also presented are

TABLE 4  
CORRELATION COEFFICIENTS: LEARNING SCORES VS. PRODUCTION, PAIRED-COMPARISON,  
AND NOBLE MEANINGFULNESS VALUES<sup>a</sup>

Scale value	Retardates		Normals	
	<i>r</i>	Rho	<i>r</i>	Rho
Noble	.80**	.89**	.95**	.95**
Production	.62*	.35	.85**	.86**
Paired-comparison	.40	.40	.71**	.82**

<sup>a</sup> These are negative correlations in that high scale values are associated with low error scores. Signs are omitted.

\*  $p < .05$ .

\*\*  $p < .01$ .

correlations between order of learning and ordinal position of scale values. In terms of statistical significance, there were no differences in the predictive validity of the various scale values, with the exception of the following comparisons involving retardate rank-order correlations: Noble *m'* vs. production ( $z = 1.96$ ,  $p < .05$ ) and Noble *m'* vs. paired-comparison ( $z = 2.20$ ,  $p < .05$ ). The Noble values predicted retardate learning more adequately than did our production and paired-comparison values. Paired-comparison values had the lowest predictive validity in all cases. The three sets of scale values appeared to predict learning success about equally well for the normal children. Correlations with learning performance were substantial for all sets of scale values. In terms of the absolute magnitudes of the correlations, however, the paired-comparison values were least successful in predicting learning scores, while Noble's *m'* had the greatest validity.

### 4. Discussion

These data yield evidence in support of the assumption that meaningfulness values of trigrams for adult and developmentally immature subjects are ordered in the same relative fashion. Furthermore, the relationship between paired-associate learning and meaningfulness (e.g., 6, 19) observed with

normal adults appears to hold true for these subjects as well. Meaningfulness values taken from adult norms were used very successfully to predict rate of learning in paired-associate tasks with children and retardates.

The results suggest that, in terms of the absolute value of correlations with learning scores, Noble's  $m'$  values have greater predictive validity than either production or paired-comparison values. When the analysis is extended to include mental retardate data, this is even more evident.

The validity of Noble's scale reflects the effectiveness of his rating method in scaling meaningfulness for adult subjects. The validity of the production and paired-comparison scale values for children and retardates reflects the effectiveness of the two techniques in scaling meaningfulness for these subjects. In order for the production and paired-comparison scale values to be useful, they should be at least as valid as Noble's values. For the normal children, the various scale values predicted learning rates about equally well, which indicated a high degree of similarity among the three distributions of scale values. Neither the production nor paired-comparison method yielded retardate scale values that greatly resembled Noble's values. In fact, the validity of Noble's  $m'$  was significantly greater than either distribution, at least in terms of ordinal position of meaningfulness value and trigram learning order. However, it should be remembered that the trigrams were originally selected on the basis of optimal separation among Noble's  $m'$  values. Moreover, it is clear that the paired-comparison method generated scale values that were both reliable and valid for predicting paired-associate learning rate.

#### REFERENCES

1. ARCHER, E. J. A re-evaluation of the meaningfulness of all possible CVC trigrams. *Psychol. Monog.*, 1960, **74**, Whole No. 497.
2. COOPER, J. C., JR., & GAETH, J. H. Interactions of modality with age and with meaningfulness in verbal learning. *J. Educ. Psychol.*, 1967, **58**, 41-44.
3. FLAVELL, J. H., BEACH, D. R., & CHINSKY, J. M. Spontaneous verbal rehearsal in a memory task as a function of age. In P. H. Mussen, J. J. Conger, & J. Kagan (Eds.), *Readings in Child Development and Personality*. New York: Harper & Row, 1970.
4. GAETH, J. H., & ALLEN, D. V. Association values for selected trigrams with children. *J. Verb. Learn. & Verb. Behav.*, 1966, **5**, 473-477.
5. GLAZE, A. J. The association value of nonsense syllables. *J. Genet. Psychol.*, 1928, **35**, 255-269.
6. GOSS, A. E., & NODINE, C. F. *Paired-Associates Learning*. New York & London: Academic Press, 1965.
7. GUILFORD, J. P. *Psychometric Methods*. New York: McGraw-Hill, 1954.
8. HULL, C. L. The meaningfulness of 320 selected nonsense syllables. *Amer. J. Psychol.*, 1933, **45**, 730-734.
9. KRUEGER, W. C. F. The relative difficulty of nonsense syllables. *J. Exper. Psychol.*, 1934, **17**, 145-153.

10. LINDQUIST, E. F. *Design and Analysis of Experiments in Psychology and Education*. Boston: Houghton Mifflin, 1953.
11. LOGAN, D. R. Paired-associate learning performance as a function of meaningfulness and response times. *Amer. J. Ment. Defic.*, 1969, **74**, 249-253.
12. MANDLER, G. Associative frequency and associative prepotency as measures of response to nonsense syllables. *Amer. J. Psychol.*, 1955, **68**, 662-665.
13. NOBLE, C. E. An analysis of meaning. *Psychol. Rev.*, 1952, **59**, 421-430.
14. ———. Measurements of association value ( $a$ ), rated associations ( $a'$ ), and scaled meaningfulness ( $m'$ ) for the 2100 CVC combinations of the English alphabet. *Psychol. Rep.*, 1961, **8**, 487-521.
15. NOBLE, C. E., STOCKWELL, F. E., & PRYER, M. W. Meaningfulness ( $m'$ ) and association value ( $a$ ) in paired-associate syllable learning. *Psychol. Rep.*, 1957, **3**, 441-452.
16. PALERMO, D. S., FLAMER, G. B., & JENKINS, J. J. Association value of responses in the paired-associates learning of children and adults. *J. Verb. Learn. & Verb Behav.*, 1964, **3**, 171-175.
17. PREHM, H. J. Associative learning in retarded and normal children as a function of task difficulty and meaningfulness. *Amer. J. Ment. Defic.*, 1966, **70**, 860-865.
18. SHAPIRO, S. S. Meaningfulness values for 52 CVCs for grade-school-aged children. *Psychon. Sci.*, 1964, **1**, 127-128.
19. UNDERWOOD, B. J., & SCHULZ, R. W. *Meaningfulness and Verbal Learning*. Philadelphia: Lippincott, 1960.

*Center for Developmental and Learning Disorders*

*University of Alabama*

*Box 6302*

*University, Alabama 35486*





## PRACTICE EFFECT IN ANAGRAM SOLVING\*

*Herbert H. Lehman College of the City University of New York*

---

EDWARD I. GAVURIN

---

### SUMMARY

An experiment was designed to study the effect of practice on anagram solving. This was accomplished through the presentation of three anagram lists to 243 subjects and a comparison made of first-list performance (early stage of practice) with last-list performance (later stage of practice). The results revealed the presence of a significantly greater number of solutions on the last, rather than the first list, clearly indicating the presence of a general practice effect in the solution of anagrams. The failure of earlier studies to demonstrate this phenomenon was attributed to the use of relatively short anagram lists in previous research.

---

### A. INTRODUCTION

The influence upon anagram solving of solution sets established by experimental instructions, word category, word sequence, letter order, or a combination of these has already been clearly demonstrated (6). Little evidence is available, however, to demonstrate the presence of a general anagram practice effect established through the solution of successive problems when solution sets are absent. In fact, there have been virtually no studies devoted exclusively to the effect of practice on anagram solving, and in the few investigations principally concerned with other variables where the data have been analyzed only secondarily for this phenomenon (1, 2, 3, 4, 5), the results on practice effects have been reported merely as auxiliary findings. In general, these studies have attempted to detect the presence of a practice effect by comparing performance on early anagram problems in a list with that of performance of later problems in the list. Although in complete agreement concerning the failure of the data to demonstrate such an effect, all of these investigations were designed to present relatively short anagram lists (from 10 to 20 problems) over short periods of time to their subjects. As a result, it is possible that the cumulative effect of practice in these experiments

---

\* Received in the Editorial Office on April 30, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

may have been below that required for detection by statistical analyses, and that longer lists were necessary to make the effect prominent. Thus, both the limitations of earlier studies, which treated the practice effect as a matter of secondary interest, and the absence of research specifically devoted to this problem suggested the need for the present investigation.

## B. METHOD

### 1. *Subjects*

The subjects were 243 undergraduate students enrolled in psychology classes at Herbert H. Lehman College.

### 2. *Materials and Procedure*

Three lists of 20 five-letter anagrams each were employed. The lists were roughly equivalent to one another with respect to solution-word frequency according to the Thorndike-Lorge word count, and were composed of problems that were selected and arranged in a way to prevent letter-order, word-category, and word-sequence sets.

The subjects were tested in several groups (classes) on all three lists, with approximately half of each group working on List 1 ( $L_1$ ), followed by List 2 ( $L_2$ ), followed by List 3 ( $L_3$ ), or sequence  $L_1L_2L_3$ , while the other half was required to follow the reverse list order (sequence  $L_3L_2L_1$ ). Consequently, performance on  $L_1$  and  $L_3$  during the early phases of practice, when they occurred in the first position, could be compared with performance on these lists during the later stages of practice, when they appeared in the last position. In this design  $L_2$ , always in the second or middle position, served exclusively as an interpolated practice session between the first and last lists.

Each anagram list was printed on a separate sheet with a space next to each problem for recording the solution. The three lists with a blank page between each list and a top page for recording identifying information were stapled together to form a test booklet. The lists in half of the booklets were arranged in the  $L_1L_2L_3$  sequence, while the other half contained the  $L_3L_2L_1$  order.

Through oral instructions, the subjects were given a description and an example of an anagram problem, told that they would be required to work on each list for exactly seven minutes in the order that these lists appeared in the test booklets, and informed that one minute of rest would intervene between lists. Following these instructions, the subjects were required to turn the top page of the booklet and to begin working on the first list. After seven min-

utes, they were told to turn to the next page (which was blank) and to rest for one minute. At the conclusion of the rest period, the subjects were instructed to turn to the next page, which contained the second list, and again were given a seven minute work period. This procedure continued until all three lists had been completed.

### C. RESULTS AND DISCUSSION

Table 1 permits a comparison of performance on  $L_1$  and  $L_3$  when each occurred as the first list in the sequence (early stage of practice) with performance on these lists when each occurred as the third list in the sequence (late stage of practice). For each of the lists, there was a significantly greater

TABLE 1  
MEANS AND STANDARD DEVIATIONS FOR THE EXPERIMENTAL LISTS SOLVED IN THE FIRST AND LAST POSITIONS

List	Position in which solved	<i>N</i>	<i>M</i>	<i>SD</i>
$L_1$	First	124	12.15	4.17
	Last	119	13.95	4.37
$L_3$	First	119	10.13	4.65
	Last	124	11.14	4.14

number of solutions when it was solved at the end rather than at the beginning of the experimental session. The  $t$  value obtained from comparing the means of List 1 was 3.28 ( $df = 241$ ,  $p < .005$ ), while the  $t$  obtained from a comparison of the means of List 3 was 1.78 ( $df = 241$ ,  $p < .05$ ). A similar comparison across both lists was made of first-list performance (combining scores on  $L_1$  and  $L_3$  when solved in the first position) and last-list performance (combining scores on  $L_1$  and  $L_3$  when solved in the last position). This revealed that the mean number of first-list solutions was 11.16 with a standard deviation of 4.53, while the corresponding values for last-list performance were  $M = 12.51$  and  $SD = 4.48$ , respectively. A correlated  $t$  test comparing these two means disclosed the presence of a significantly greater number of solutions during last-list performance ( $t = 5.52$ ,  $df = 242$ ,  $p < .0005$ ).

In view of the fact that for all of the above comparisons, a greater number of solutions occurred at the end rather than at the beginning of the experiment, the presence of a general practice effect seems to be reliably established. This clear-cut effect of practice, undetected in earlier studies, was achieved in the present investigation through the use of relatively long problem lists which

provided for sufficient experience in the solution of earlier problems to facilitate the solution of later ones. The failure of previous studies to produce any evidence of this phenomenon can probably be attributed to the use of relatively short anagram lists which did not provide the minimal conditions necessary for a measurable change in performance.

#### D. CONCLUSION

From the results of the present study, it is possible to conclude that anagram solving skill clearly improves with practice, and that demonstrable improvement requires the use of relatively long practice lists. Future studies should consider the functional relationship between the degree of practice and the extent of improvement.

#### REFERENCES

1. BEILIN, H., & HORN, R. Transition probability effects in anagram problem solving. *J. Exper. Psychol.*, 1962, **63**, 514-518.
2. DOMINOWSKI, R. L., & DUNCAN, C. P. Anagram solving as a function of bigram frequency. *J. Verb. Learn. & Verb. Behav.*, 1964, **3**, 321-325.
3. HUNTER, I. M. L. The solving of five-letter anagram problems. *Brit. J. Psychol.*, 1959, **50**, 193-206.
4. JENKINS, H. M., & MAYZNER, M. S. Time-pattern of switching among alternatives in problem solving. *Psychol. Rep.*, 1961, **8**, 419-430.
5. MAYZNER, M. S., & TRESSELT, M. E. Anagram solution times: A function of word transition probabilities. *J. Exper. Psychol.*, 1962, **63**, 510-513.
6. REES, H. J., & ISRAEL, H. E. An investigation of the establishment and operation of mental sets. *Psychol. Monog.*, 1935, **46**, No. 210.

*Department of Psychology*  
*Herbert H. Lehman College of the*  
*City University of New York*  
*Bronx, New York 10468*

## THE EFFECTS OF VERBAL REPORTS OF VIOLENCE ON AGGRESSION\*<sup>1</sup>

*Oklahoma State University*

---

WILLIAM H. SCHARFF<sup>2</sup> AND ROBERT S. SCHLOTTMANN<sup>3</sup>

---

### SUMMARY

Forty male and 40 female volunteer college students were divided into two groups. Half were exposed to actual tape recorded verbal reports of violent events, and half to similar reports of nonviolent happenings. Ss who had been angered by insult prior to being exposed to violent tapes displayed significantly more aggression than Ss in an insult, nonviolent condition and Ss in a no-insult, violent condition on a subsequent "extrasensory learning" task supposedly involving shocks for incorrect responses. An unexpected finding was that Ss who had not been insulted administered significantly higher "shocks" after exposure to nonviolent reports than Ss in the no-insult, violent group. No significant sex differences were found. Results were interpreted as failing to support the catharsis hypothesis, and comparisons with the effects of visually witnessed violence were made.

---

### A. INTRODUCTION

Klapper (10) reported that a considerable number of psychiatrists and psychologists interviewed by him in 1953 thought that the display of aggression in the communications media provided an opportunity for hostility catharsis. The catharsis hypothesis of hostility contends that the performance of an aggressive act reduces the tendency to act out aggression and does not necessarily involve elimination of the stimulus that elicited the aggression.

---

\* Received in the Editorial Office on May 1, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This paper is based upon a thesis submitted by the first author to Oklahoma State University in partial fulfillment of the requirement for the M. S. degree. The authors thank Drs. Kenneth Sandvold and Barbara Weiner for their advice and assistance as committee members, and Earl Chiddix, Dave Ferguson, Cecil Norris, and Lyn Thomas who participated in the collection of data.

<sup>2</sup> Now at the Diagnostic and Classification Center, Cornwells Heights Youth Development Center, Cornwells Heights, Pennsylvania.

<sup>3</sup> Requests for reprints should be sent to the second author at the address shown at the end of this article.



Support for the catharsis hypothesis had previously been obtained by Feshbach (8). An experimental group was first angered by insult and then given four cards of the Thematic Apperception Test (TAT) which served as a measure of aggression that the insult induced and as a means of catharsis. The remaining two groups were used as controls; one was angered but did not take the TAT, and the other not angered but took the TAT. Ss who were insulted produced significantly higher aggression scores on the TAT than those who were not insulted. A questionnaire involving attitudes toward the experimenter and the experiment demonstrated that the insult-catharsis group displayed significantly less aggression than the insult-control group. Thus, the telling of aggressive stories in the TAT lowered the tendency of angered Ss to aggress.

A number of studies have more recently questioned the cathartic value of violence in the mass media. Experiments using television programs (7), cartoons (13, 14), and film strips (1, 4, 16) have consistently shown that witnessed violence can actually increase the probability that the viewer will act aggressively. Berkowitz (3) explains these as possibly due to the acquiring of new aggressive actions by imitating the aggressor on the screen. Also, exposure to violence can make aggression morally proper and may lower inhibitions against aggression in the viewer (5).

The above studies seem to demonstrate that the viewing of violence increases the probability of further aggression. The present study investigated the possibility that violence need not be witnessed visually to elicit aggression. It seemed possible that simply hearing actual verbal reports of nonfictional violent occurrences might also increase the probability of aggressive behavior in previously angered Ss. Without the aid of pictures, verbal reports may involve very descriptive audio cues leaving more to the imagination than the narrower film strip in which many aspects of the situation are predetermined. Actual, nonfictional reports were chosen because of their greater cue similarity to real life situations. Also, their high frequency of occurrence as reported in the mass media may have a desensitizing effect and serve to inhibit the anxiety from aggression.

## B. METHOD

### 1. *Subjects*

Forty male and 40 female freshmen and sophomore college students from introductory psychology classes at Oklahoma State University were used. All Ss volunteered for the experiment for extra course credit.

## 2. Apparatus

The apparatus was similar to that used by Buss (6). It consisted of a  $12.5'' \times 24.5'' \times 12.5''$  board containing a series of 10 levers numbered in order from 1 to 10. The words "mild" and "strong" were written just beneath levers 1 and 10, respectively. Another lever was located in the middle of the board beneath the 10 levers and designated as "ready." Levers were connected to a series of lights in an adjacent room numbered from 1 to 10 to correspond to each of the 10 levers. An additional light, designated as "alert," corresponded to the "ready" lever.

## 3. Procedure

Each *S* was tested separately with a confederate who passed as a student from a different class. *Ss* were told that they would be participating in a series of learning tasks. The first task of the *S* was to shock the confederate for incorrect answers on a supposedly "extrasensory learning" experiment. No shock was actually given, since levers on the board were connected only to a series of lights in the adjacent room to record the level of "shock" administered. A list of 12 colors was presented to *S*. The task of *S* was to press the alert lever which signaled the confederate when *S* was thinking of a color in the predetermined order. The confederate had a list of colors from which he responded that contained two correct (matching) colors. *S* was instructed to press the alert lever when thinking of a color on his list. Going in order down the list was stressed. *S* was further instructed to give any degree shock desired when a wrong answer was given over a microphone by the other *S* (confederate) in the adjacent room. He was also told that the shocks ranged from mild (#1) to strong (#10), but were not strong enough to hurt anyone seriously. *E* then left the room leaving *S* alone to choose any shock desired for incorrect responses.

Once the list was completed, *E* and the confederate returned to *S*'s room where the second task, which was a multiple-choice number test, was given. This task was explained as a learning-type exercise in which both *Ss* (the *S* and confederate) would be thinking of the same things. The test consisted of a series of four numbers in which *S* had to pick the correct reversed order out of four possible alternatives. There were 20 series presented. A time limit of one minute was set making it impossible to complete the test with accuracy. *E* remained in the room during the task.

Half of the *Ss* (20 males and 20 females) took this test without incident. The remaining *Ss* were insulted by the confederate who completed the test

in the time allowed (since he already knew the correct answers). The confederate (a man in his early twenties) complained to *E* that the test was too simple, asked *S* if he (she) had finished, and sarcastically chided *S* for having failed to complete it. The procedure was not memorized but held flexible to the answers of *S*.

In the third task, *S* and the confederate were asked to relax and listen to portions of news broadcasts spliced together on a tape recorder. This task was explained as a learning task involving current events. Twenty male and 20 female *Ss* listened to a tape with violent news reports (i.e., stabbings, a bombing, shootings, and a riot). Of this group, 10 males and 10 females were of the insult group, and 10 males and 10 females were from the no-insult group. The remaining *Ss*, 20 from the insult and 20 from the no-insult group, listened to relatively calm, nonviolent broadcasts (i.e., due date of license plates, a turkey call contest, and an election of an Indian chief). Both tapes were approximately five minutes in length. *Ss* were told that some questions would be asked at the end of the broadcasts. There were three multiple-choice questions for each tape to make sure *Ss* attended to the tapes.

After listening to the tape (either violent or nonviolent), *S* and the confederate were again asked to participate in the "extrasensory learning experiment," ostensibly to see if their thinking together on the other tasks had enhanced the ability to perceive each other's thoughts. At the end of the experiment, *E* explained the actual purpose of the study to *S*.

The general design for this experiment was a  $2 \times 2 \times 2$  analysis of covariance. The three factors under study were the insult condition (insult *versus* no-insult), type of tape (violent *versus* nonviolent), and sex of *S*. For each *S*, two scores were recorded which consisted of the means of the pre- and posttest shock task. The mean pretest was used as the covariate for each *S*, and the mean posttest the dependent variable in the analysis.

### C. RESULTS

The means and standard deviations of pretest and posttest shock levels administered are shown in Table 1 for *Ss* in each group. Analysis of covariance revealed a significant interaction between the insult (no-insult) and violent (nonviolent) tape conditions ( $F = 14.88$ ,  $df = 1, 71$ ,  $p < .01$ ). Figure 1 shows the adjusted mean shock levels administered by insulted and non-insulted *Ss* after exposure to violent or nonviolent tapes. Investigation of the simple main effects for the interaction of the adjusted means showed that insulted *Ss* who were exposed to tapes of verbal reports of violence gave

TABLE 1

MEANS AND STANDARD DEVIATIONS OF PRETEST AND POSTTEST SHOCK LEVELS ADMINISTERED

Group	Pretest		Posttest		Difference
	Mean	SD	Mean	SD	
Insult, violent tape					
Male	5.26	1.23	6.63	1.38	+1.37
Female	5.12	1.83	6.96	2.00	+1.84
Insult, nonviolent tape					
Male	4.14	1.94	4.42	1.87	+ .28
Female	4.14	1.34	4.61	2.01	+ .47
No-insult, violent tape					
Male	5.24	2.25	5.25	2.20	+ .01
Female	4.24	1.20	4.12	1.78	— .12
No-insult, nonviolent tape					
Male	5.14	1.01	6.40	1.25	+1.26
Female	4.27	1.73	4.58	1.80	+ .31

significantly higher shocks than insulted Ss who were exposed to tapes involving verbal reports of nonviolent happenings ( $t = 3.362$ ,  $df = 71$ ,  $p < .05$ ). It was also found that insulted Ss gave significantly higher shocks than noninsulted Ss after exposure to verbal reports of violence ( $t = 4.255$ ,  $df = 71$ ,  $p < .05$ ). Lastly, noninsulted Ss gave significantly higher shocks after exposure to verbal reports of nonviolence than after exposure to verbal reports of violence ( $t = 2.067$ ,  $df = 71$ ,  $p < .05$ ).

#### D. DISCUSSION

The results of the present study are in general agreement with other studies involving visually witnessed violence in failing to support the catharsis hypothesis. Exposure to verbal reports of violence can also lead to an increase in aggressive behavior in angered persons and does not serve as an outlet through vicarious satisfaction for hostile tendencies. Insulted Ss, after being exposed to verbal reports of stabbings, a bombing, shootings, and a riot, administered significantly higher shocks than insulted Ss who were exposed to tapes containing news of a turkey call contest, due date of license plates, and other more innocuous material.

But exposure to violence by itself was not sufficient to elicit aggression. The results of the present study are also in agreement with Berkowitz (2) in suggesting that a predisposition to behave aggressively is necessary. As indicated by analysis of simple main effects for the interaction of the adjusted means, insult alone and exposure to reports of violence alone had relatively little effect on the level of shock administered. However, when the two conditions were both present (i.e., that insult followed by exposure to taped reports of violence), the shock level increased significantly. It appears, as

Berkowitz and Rawlings (5) suggested, that exposure to violence will tend to lessen an angered listener's inhibitions against aggression. The fact that significantly lower shocks were administered by Ss who were insulted and exposed to reports of a more calm nature seems to indicate that these Ss'

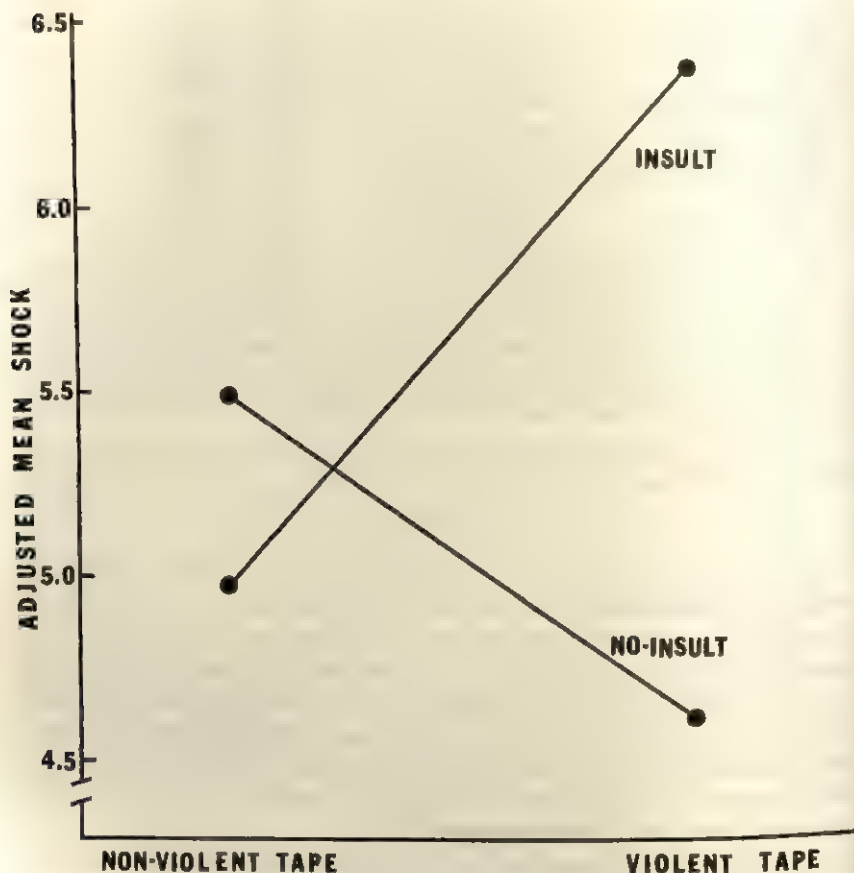


FIGURE 1

ADJUSTED MEAN SHOCKS FOR INSULTED AND NONINSULTED SUBJECTS IN VIOLENT AND NONVIOLENT TAPE CONDITIONS

inhibitions were still in effect. In fact, during informal posttest questioning and debriefing, many Ss in this group expressed contempt for the insulter, but reported they were able to control their anger.

An interesting finding in the present study is that female Ss were just as



aggressive as male Ss. Studies of aggression using American children and teenagers as Ss suggest that males tend to be more aggressive. Jersild and Markey (9) observed 1500 conflicts among 54 nursery school children in free play and found that girls were more likely to use verbal attacks and boys to use physical means. Levin and Sears (12) observed children's doll play and reported boys to be more aggressive than girls. Sears (15) studied aggression in 12-year-olds and found that boys scored higher on a self-report measure of antisocial aggression than girls. Lansky, Crandall, Kagan, and Baker (11) obtained comparable results using teenagers. Berkowitz (2, p. 269) states that aggressive habits are stronger in males, and males are thus likely to make stronger hostile responses to an evoking cue than females.

It seems that the display of aggression is generally expected of the male and is to an extent socially approved, whereas females are expected to inhibit their aggression. It may be that the exposure to actual verbal reports of factual events as used in the present study was "powerful enough" to break through the females' stronger inhibitions. There is also some question as to whether similar results would be found if the insulter were a woman. Since all insulters in the present study were males, possible interaction effects between sex of insulter and sex of *S* were not investigated. It seems plausible that such effects could be due to possible differences in the stimulus value of same sex or opposite sex insulters in arousal potential and/or ability to break through inhibitions.

A totally unexpected finding was that Ss who were not insulted displayed significantly more aggression after hearing taped reports of nonviolent occurrences than those who were likewise not insulted but later exposed to taped reports of violent incidents. At first, this finding may appear to support the catharsis hypothesis. However, the catharsis model involves displacement of aggression. In the no-insult condition, no anger was instilled in Ss. Informal posttest questioning revealed that many Ss in the no-insult, non-violent tape conditions were annoyed over failure of the confederate to demonstrate improvement on the "extrasensory learning" task and, therefore, increased the level of shock administered for incorrect responses. Such annoyance did not seem as pronounced in the no-insult, violent tape condition, suggesting the possibility that these Ss may have been distracted from the "extrasensory learning" task by exposure to the actual reports of violence. In addition, the catharsis model involves a reduction in the tendency to aggress after exposure to violence. As can be seen in Table 1, Ss who were not insulted but later exposed to taped reports of violent events administered about the same level of shocks on the posttest as they did on the pretest.

## REFERENCES

1. BANDURA, A. D., ROSS, D., & ROSS, S. A. Imitation of film-mediated aggressive models. *J. Abn. & Soc. Psychol.*, 1963, **66**, 3-11.
2. BERKOWITZ, L. Aggression: A Social Psychological Analysis. New York: McGraw-Hill, 1962.
3. ———. Experimental investigations of hostility catharsis. *J. Consult. & Clin. Psychol.*, 1970, **35**, 1-7.
4. BERKOWITZ, L., & GREEN, R. G. Film violence and the cue properties of available targets. *J. Personal. & Soc. Psychol.*, 1966, **3**, 525-553.
5. BERKOWITZ, L., & RAWLINGS, E. Effects of film violence on inhibitions against subsequent aggression. *J. Abn. & Soc. Psychol.*, 1963, **66**, 405-412.
6. BUSS, A. H. The Psychology of Aggression. New York: Wiley, 1961.
7. ERON, L. D. Relationship of T.V. viewing habits and aggressive behavior in children. *J. Abn. & Soc. Psychol.*, 1963, **67**, 193-196.
8. FESHBACH, S. The drive-reducing function of fantasy behavior. *J. Abn. & Soc. Psychol.*, 1955, **50**, 3-11.
9. JERSILD, A. T., & MARKEY, F. V. Conflicts between preschool children. *Child Devel. Monog.*, 1935, No. 21.
10. KLAPPER, J. The Effects of Mass Communication. New York: Free Press, 1960.
11. LANSKY, L. M., CRANDALL, V. J., KAGAN, J., & BAKER, C. T. Sex differences in aggression and its correlates in middle class adolescents. *Child Devel.*, 1961, **32**, 45-58.
12. LEVIN, H., & SEARS, R. R. Identification with parents as a determinant of doll-play aggression. *Child Devel.*, 1956, **27**, 135-153.
13. LOVAAS, O. I. Effect of exposure to symbolic aggression on aggressive behavior. *Child Devel.*, 1961, **32**, 37-44.
14. MUSSEN, P., & RUTHERFORD, E. Effects of aggressive cartoons on children's aggressive play. *J. Abn. & Soc. Psychol.*, 1961, **62**, 461-464.
15. SEARS, R. R. Relation of early socialization experiences to aggression in middle childhood. *J. Abn. & Soc. Psychol.*, 1961, **63**, 461-465.
16. WALTERS, R. H., THOMAS, E. L., & ACKER, C. Enhancement of punitive behavior by audio-visual displays. *Science*, 1962, **136**, 872-873.

*Department of Psychology*  
*Oklahoma State University*  
*Stillwater, Oklahoma 74074*

## STUDIES IN THE NEUROPHYSIOLOGY OF LEARNING: VIII. OSCILLATORY POTENTIALS RESULTING FROM CEREBRAL SELF-STIMULATION IN RATS\*<sup>1</sup>

*University of California, Los Angeles*

---

J. A. GINGERELLI

---

### SUMMARY

1. Evoked potentials generated by cerebral self-stimulation were studied in white rats. Stimulus trains of 30 biphasic rectangular waves of various repetition rates (30, 50, 100, 150, 200 pps) were employed. The pulses were always .4 msec in width and of fixed intensity at all repetition rates for a given animal. Potential changes during a 2-second interval after the termination of a stimulus train were recorded from various points in the cerebral mass and averaged electronically.

2. The early portion of the 2-second interval was characterized by rhythmic oscillations in the alpha range whose number and amplitude were related to repetition rate of the rectangular waves within the stimulus train. The periodicity of the oscillations when they occurred was not so affected and remained practically constant over subjects, intensity of stimulus, and repetition rate.

3. The number of oscillations was greatest at 200 and 150 pps and diminished with diminishing stimulus repetition rates, the drop being greatest below 100 pps.

4. There was a positive covariation between the rate of lever pressing for self-stimulation and the phenomenon of rhythmic after oscillations in the evoked potential. Animals who refused to press the lever for cerebral self-stimulation showed no rhythmic potentials when the stimulus was delivered by *E*; animals who had become pressers gave good after-discharges when so stimulated.

---

\* Received in the Editorial Office on May 2, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This work was supported by the UCLA Faculty Research Committee and expedited by an item of equipment donated by the Holter Research Foundation, Helena, Montana. The author wishes to express his sincere thanks to Mr. Lon Davis, electronics technician, and to Mr. Darrell Dearmore, who prepared the illustrative material, for their valuable assistance.

---

### A. GENERAL

The phenomenon of reinforcement is the central mystery in the learning process. This is very probably the case irrespective of the level at which the learning problem is considered; whether the approach is phenomenological or neurophysiological, we need to understand why it is that certain events are "positively reinforcing" in the sense that they are conducive to increased probability of a defined reaction, whereas others result in a diminished probability. More concretely, what is it that positive reinforcers have in common; what is it that negative reinforcers have in common? There must be something common to members of each class if we are to construct a theory worthy of the name.

In the context of neurophysiological theory, the problem, expressed in its most general terms, takes the form of understanding why the reaction to a fixed pattern of stimulation may change, for example, in one direction rather than its opposite as a consequence of the pattern's having impinged upon the central nervous system at least once. Further, it is required that any independent variables considered for this purpose be restricted to the neurophysiological domain: for instance, what parameters of neurone (or glia) functioning provide the residual effects that cause the "tide" of neural events to change in one direction or its opposite as a result of past stimulation?

It would appear that the phenomenon of cerebral self-stimulation offers an encouraging approach to the problem of positive and negative reinforcement, and the vast literature that has accumulated in a relatively short time on this subject attests to its importance for theorists and experimentalists alike. In this phenomenon we can control the locus and the parameters of stimulation, and record the consequences accurately insofar as behavior is concerned. What is required, however, is more information on the cerebral electrographic accompaniments of "successful" and "unsuccessful" stimulation. This is what has been attempted in the present investigation. The procedure is frankly exploratory, with no attempt at histological determination at the present time. Histology can be quite unfruitful for neurological learning theory in the same sense that circuit diagrams are not very revealing without knowledge of the fundamental laws of electricity and magnetism and the characteristics of the component parts. The basic processes underlying neural plasticity will be found very probably at the molecular rather than the morphological level.

## B. METHOD

1. *Electrode Assembly and Placement*

The electrode assembly is shown in Figure 1, where the distances between the various electrodes and their configuration are given. The block was constituted of three layers of Amphenol Strip Connectors cut to the desired length and cemented together. Stainless steel wires, .017" in diameter, were crimped into the bottom of the male contact pins (Amphenol) and inserted into the holes provided by the strips. The wires were then cut to the desired length, and constituted the electrodes, insulated except at the tip. The resulting structure was a block 11 mm high, 10 mm long, and 7 mm wide, exclusive of the electrodes (*cf.* Figure 1 for details).

Holes were made in the rat's skull by inserting a small burr drill through holes in a jig that copied the configuration of the electrodes emerging from the block; before drilling began, the jig was fitted to the stereotaxic instru-

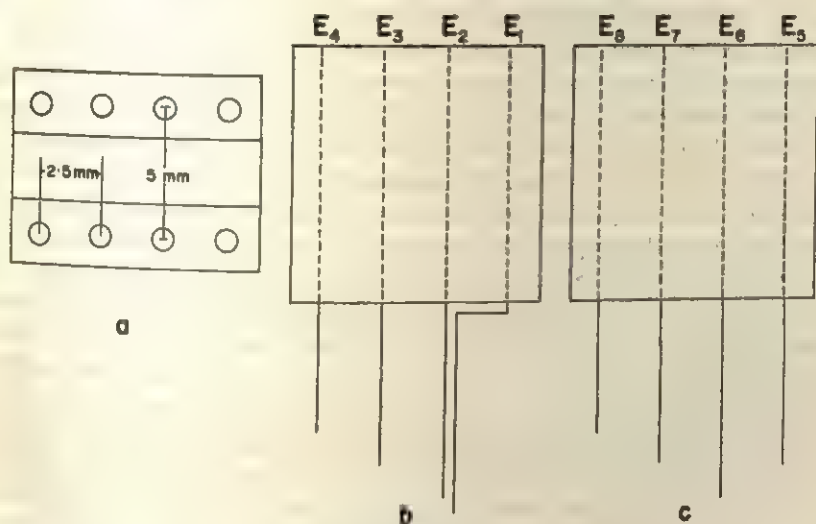


FIGURE 1

## DIAGRAMMATIC REPRESENTATION OF ELECTRODE UNIT

(a) Top View; shows the three amphenol strips, with electrodes placed in the two outside strips. (b) Electrode arrangement in the right lateral strip, with electrodes numbered; E<sub>1</sub> and E<sub>2</sub> are the stimulating electrodes. (c) Electrode arrangements in the left lateral strip. The unit is so oriented that the right end is forward: E<sub>1</sub> = 8.5 mm; E<sub>2</sub> = 8 mm; E<sub>3</sub> = E<sub>7</sub> = 6 mm; E<sub>4</sub> = E<sub>8</sub> = 3-4 mm; E<sub>5</sub> = 7 mm; E<sub>6</sub> = 8 mm.



ment and lowered carefully upon the rat's skull. The orientation of the jig was such that the right-hand row of holes fell uniformly about 2 mm to the right of the skull midline, and the "lead hole"—hole #2—was 1.5 to 2.0 mm behind bregma. This "fix" automatically determined the loci of all other holes. The electrode assembly was then affixed to the stereotaxic and lowered carefully in such a way that each electrode was inserted into its intended hole (a surprisingly demanding task!).

Electrical contact with the electrodes was made through a corresponding block unit fitted with female contact pins (Amphenol). The coupling between the pick-up block and the leads to the amplifying instruments was by means of Zenith hearing aid wires, 43" long, eight of which were carefully braided together. This formed a very light "cable" of extraordinary flexibility and minimally plagued, as far as I could determine, by intruding noise.

## *2. Apparatus, Circuitry, and Behavior Box*

The stimuli were furnished by two Grass SD5 stimulators, connected in cascade, such that one instrument furnished the positive phase of the biphasic stimulus and the other, the negative. The stimuli, unless otherwise indicated, were applied to the animal through electrodes 1-2. Any two of the other six electrodes could serve as pick-up units. The amplifying instruments were two Tektronix Type 122 preamplifiers, coupled in cascade and set at a combined amplification of 10,000; range of frequencies set at .8 cycles/second to 50 cycles/second. In any given experiment the signals from two electrodes were amplified, while the output from the other four "idle" electrodes was coupled to the ground (reference) input of the amplifier. Any desired combination of "recording" and "idle" electrodes was accomplished by means of a switching panel whereby the two recording electrodes were coupled to the "hot" pins of the amplifier, and the four idle ones were coupled to a metal bar feeding into the ground pin. (See Figure 2.) The amplified output from the recording electrodes was fed into a Hewlett-Packard averaging apparatus (5480A).

Recording from a brain that has very recently been stimulated by an extended train of impulses presents certain technical difficulties. The brain tissue and the assemblage of electrodes and conductors form a capacitative system which, after stimulation is ended, discharges over a period of time that intrudes upon physiological events to be recorded. The resulting artifact, recorded by the pick-up electrodes, is capable of distorting the record to an undesirable degree.

The circuitry that was found effective in mitigating very strong artifacts residual to stimulation is shown in block fashion in Figure 2. A 5 megohm

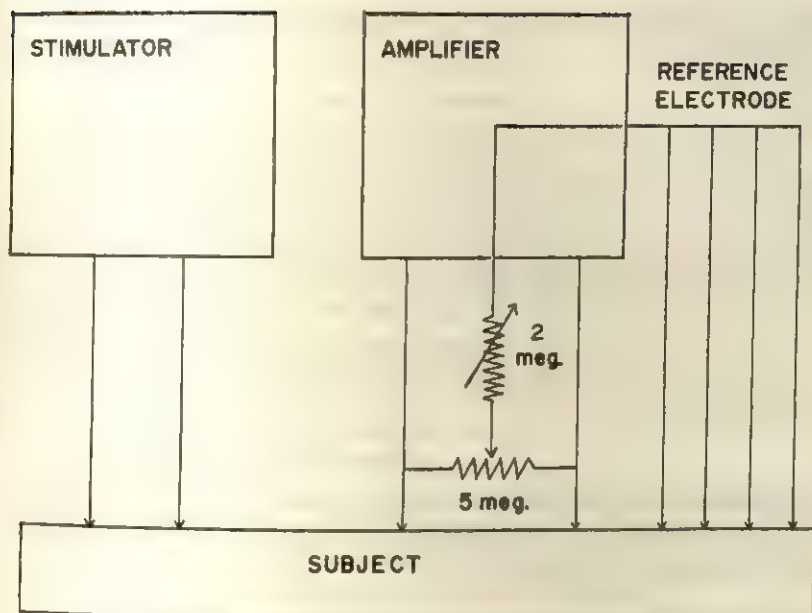


FIGURE 2

DIAGRAMMATIC REPRESENTATION OF STIMULATION AND PICK-UP CIRCUITRY

potentiometer placed across the leads from the recording electrodes is tapped at various points, and the output coupled to the reference (ground) bar through a 2 megohm variable resistor. Manipulation of these two variables, as dictated by the size and pattern of the registered artifact created by the stimulus train, served to "bleed off" and diminish the charge impinging upon the pick-up electrodes. The settings made to minimize the stimulus artifact did not seem to affect in any gross fashion the output from the recording electrodes.

The rat was placed into a mental cylinder 15" in diameter and 22" tall. This cylinder was inside a larger metal cylinder and electrically insulated from it. The bottom of the inner cylinder was provided with an elevated metallic mesh to avoid the accumulation of urine and fecal matter. The larger cylinder was coupled to earth. The rat was thus grounded only through the four idle electrodes by means of the common bar feeding into the ground pin of the amplifier. The behavior chamber was fitted with a removable lever which, when pressed, yielded a preset stimulus train. Otherwise, the animal

could be stimulated by *E*. The rat's behavior was visible at all times by a positioned mirror.

### 3. *Stimulus Parameters*

Stimuli were biphasic rectangular pulses, each phase .2 msec in width and, unless otherwise specified, delivered in trains of 30 pulses at the following repetition rates: 30, 50, 100, 150, 200 pps. This, of course, altered the duration of the train for the various rates. While the stimulation was being applied, the stimulus impulse configuration was monitored oscilloscopically at all times. Circuitry was provided to trigger the sweep of the averager at the end of each stimulus train. In one experiment (as indicated) the sweep was triggered at the onset of the stimulus train. The sweep used was usually of 2 seconds duration.

### 4. *Animals*

The animals were male albino rats, about 2 years old and weighed slightly in excess of 400 grams. After implantation of electrodes, 10 days elapsed before experimentation began.

### 5. *Shaping and General Experimental Procedure*

The rat was first placed into the behavior chamber in the absence of the lever, and the stimulator set at a repetition rate of 150 or 200 pps. The voltage was very slowly increased while his behavior was observed. At the point when application of the stimulus train elicited alertness and "interest," voltage was increased no further, or only slightly, and the rat was stimulated by *E* at irregular intervals. If the 1-2 electrodes had "hit the target," his behavior became increasingly energetic upon stimulation and seemed directed "to finding a source." At this point the lever was introduced into the cylinder, and the shaping process began. One or two sessions were devoted to developing a good and consistent "presser."

When this was accomplished, the experiment proper was begun for that animal. The first portion of the session was devoted to self-stimulation with the 30-pulse train at 200 pps. This lasted until 128 sweeps were elicited; sweeps were elicitable only at intervals equal to 2 seconds plus the duration of the stimulus train, and the presses of the lever more frequent than this resulted in no stimulation. The animal was then taken out of the experimental compartment and rested for 5 minutes, after which he was returned, with the repetition rate set at the next lower value (150 pps), but, of course, at the

same voltage. He pressed for approximately 1 minute at the new rate before recordings were made. This pattern was observed until all five repetition rates had been applied.

When the repetition rate reached 50 pps or when the stimulating electrodes were "off target," the animals pressed sporadically, at long intervals, or not at all, thus forcing *E* to stimulate them by hand in order not to prolong the session indefinitely. As far as could be ascertained, there were no essential differences between averaged tracings resulting from hand stimulation and from self-stimulation (see Results). These comparisons were made with repetition rates of 200 and 150 pps.

The evoked potentials of both presser and nonpresser rats were recorded from various (but not all) pairs of electrodes (excluding 1-2), but since histological verification of electrode placement was not made, there will be no attempt to analyze these data by pairs.

In presenting the results, it is sometimes necessary to indicate both the stimulating and the pick-up electrodes for purposes of comparison. The following abbreviations will be used:  $St_{i,j}$ ,  $P_{k,l}$  where  $i-j$  represent the stimulating electrodes, and  $k-l$  the pick-up (active, recording) electrodes.

A factorial experimental design in applying the various repetition rates is preferable to the one used, but since there were no evident fatigue effects, and since the data were so uniform and consistent, it was felt that a burdensome statistical pattern could be dispensed with.

## C. RESULTS

### 1. General Behavioral Observations

Although reported incidentally by many investigators of cerebral self-stimulation, it is relevant to dwell on certain characteristics of rat behavior when a positive reinforcement region in the brain is being optimally stimulated. The behavior is most clearly observable when there is no lever in the apparatus, and the animal is being stimulated by *E*. If the "target" has been hit by the electrodes, the effect of optimal stimulation is apparent on the very first "shot." The rat shows an immediate increase in alertness, moves quickly about, sniffing the floor and walls of the container. The *E* is then practically certain that the placement was successful and that *S* will be a good "presser." This pattern of behavior is enhanced and becomes more dramatic as the stimulation is repeated. Characteristically, the rat sniffs violently over the entire floor of the apparatus and frequently stands up on



his hind legs. However, in my experience, it is the floor that engages his efforts primarily. His dynamics are those of an animal seeking something of overwhelming interest hidden in the floor of the container.

As aspect of the rat's behavior that merits special attention may be described as an increasing addiction to the stimulation. Whereas on the first day the animal's reaction is quite unmistakable, his reaction to priming stimulation on succeeding days may be described as frantic, though the parameters of the stimulus have not been changed; he darts about sniffing violently as if chasing something that eludes him. This seeking behavior becomes so prepotent with practice that even a rat who has become an experienced presser may ignore the presence of the lever at the beginning of an experimental session and embark on a bout of violent floor-sniffing in response to a stimulus administered by *E*.

By contrast, if the electrodes are "off the mark" or if the stimulus is not optimal, stimulation results in increasing quiescence—unless, of course, the voltage is too high, in which case there is violent somatic reaction and perhaps squealing. In many such instances, with increasing numbers of hand stimulations, some rats become completely motionless, as if asleep.

## 2. Quantitative Data

*a. Repetitive after-potentials and stimulus repetition rates.* The essential features of the evoked potentials resulting from four of the five repetition rates used are exhibited by the tracings in Figure 3. In this Figure the trace for 30 pps is not given. The records shown are characteristic of all those that have been made in this investigation: there is a definite covariation between the value of stimulus repetition rates used and the magnitude and number of after-potentials. For example, at 30 pps we have never observed a repetitive potential. The duration of the trains of repetitive potentials showed a range among various animals, under optimal stimulus repetition rates (200 or 150 pps), of .5 to 1.8 seconds. However, since no deliberate attempt was made to determine the exact voltage threshold for each animal, this range cannot be considered conclusive. A feature of the periodic phenomenon is the decline in amplitude with time. The period of the waves was found to vary from 120 to 160 msec (8-6 cycles/second), with a tendency in some cases for the period to be slightly longer in the late portions of the record. On the whole, however, the periods remained quite constant and were not influenced by repetition rate. As the repetition rate of the stimulus (voltage and other parameters constant) diminished from 200 pps, the duration of the rhythmicity diminished as did the amplitude of the waves. The



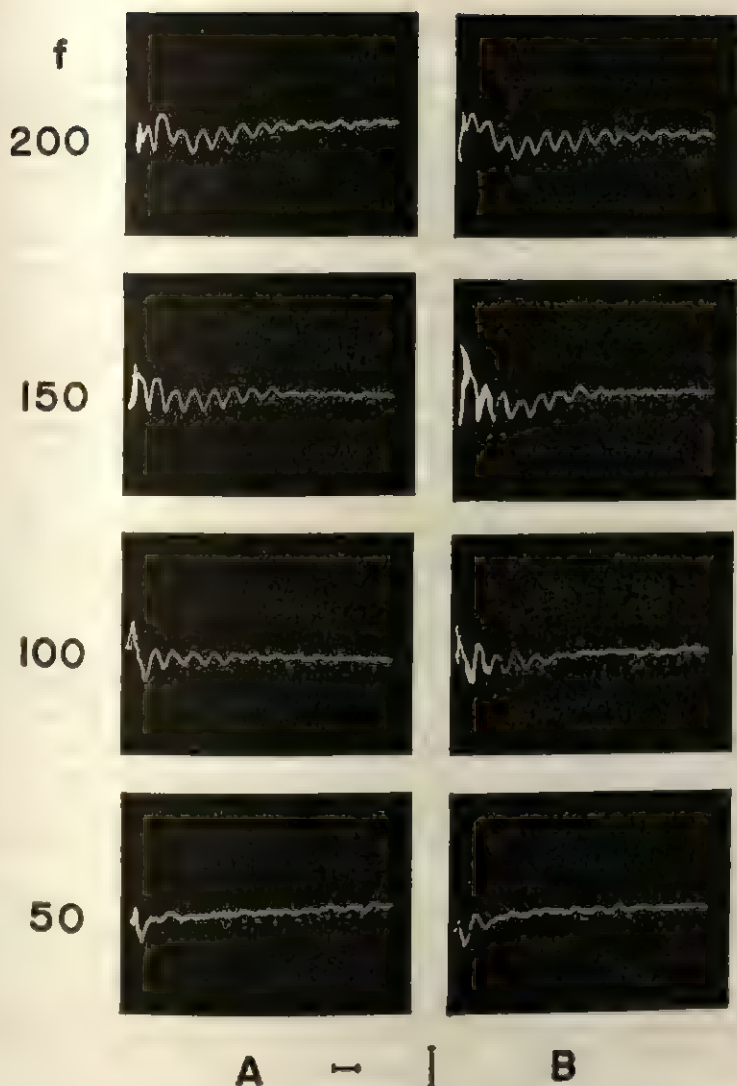


FIGURE 3

CHARACTERISTIC TRACINGS FROM THE SAME ANIMAL, SHOWING RELATION BETWEEN AFTER-DISCHARGES AND REPETITION RATE OF STIMULUS WITHIN THE GIVEN STIMULUS TRAIN

Repetition rates are given on the left side of the picture. Column A, with number of stimuli in train constant (30 pulses) for each repetition rate; column B, with duration of stimulus train constant (.2 second) for each repetition rate. The tracings here are each based on 256 sweeps; duration of sweep, 2". For parameters of stimulus, see text.

Unless otherwise indicated, in this and subsequent pictures the sweep begins at end of stimulus train; the first 80-100 msec of tracings show the effects of residual artifact of stimulus train.

Self-stimulation. Calibrations: in this and all other photographs, vertical line = 35  $\mu$ V; horizontal line = 200 msec.

biggest "break" in this respect occurred at repetition rates below 150 pps. The amplitude of the largest waves, those nearest the termination of the stimulus train, varied under optimal conditions between 20 and 80  $\mu$ V but the value of the amplitude is difficult to estimate, since it diminished as the averaged number of sweeps increased.

The tracings in Figure 3 were generated as the rat pressed the lever for self-stimulation; analogous curves have been obtained in the absence of the lever, with stimulations delivered at random intervals by *E*. The only difference noted in some instances was a smaller deflection in the record immediately after the termination of the stimulus train in the latter case. It should be pointed out that in Figure 3, as in others when repetition rates of 50 and 30 are used, all records involving 50 and 30 cps were obtained by hand stimulation of the animal, since at these frequencies he would press only in a desultory fashion.

Some experiments were performed in which the duration of the stimulus train was made constant at .2 second for all repetition rates. The effect of this, of course, was to alter the number of pulses per stimulus train: 40 at 200 pps, 30 at 150, 20 at 100, and 10 at 50. The characteristic resulting evoked potentials are shown in Figure 3B. The differences between Figure 3A and 3B—i.e., between fixed number of pulses per train (30 pulses) and fixed train duration (.2 second)—are not very impressive, although the latter seem to be "weaker" at and below 100 pps. It is noteworthy that the two (A and B) 150 pps traces involve identical stimulus parameters but are not identical, though very similar; B was taken seven days after A. All traces in the Figure are taken from the same animal and, of course, from the same electrodes.

*b. Rate of lever pressing as related to stimulus repetition rate.* Perhaps the most striking feature of the present and earlier experiments (6) is the relation between repetition rate within the intracranial stimulus train and the rate of lever pressing by the rat to obtain it. Here we repeat work previously reported (6), utilizing different stimulus parameters. In one experiment, 30 pulses were given at the five repetition rates: 200, 150, 100, 50, 30 pps. The results are shown in Table 1, where the total number of presses registered by an electric counter in a 5-minute period are given. The circuitry was such that one unit was registered for each press, regardless of frequency of pressing.

It is clear, however, that these "raw scores" present a certain ambiguity. Since the duration of the stimulus trains varied with stimulus repetition rate, it is not unreasonable to suppose that an animal might not press during

TABLE 1

NUMBER OF PRESSES IN 5-MINUTE PERIOD AS RELATED TO STIMULUS REPETITION RATE: NUMBER OF PULSES PER TRAIN, CONSTANT (30 PULSES); TRAIN DURATION, VARIABLE

PPS	# 740	# 745	# 734	# 735	# 742	# 750	# 759	# 762	# 763	N	$\bar{X}$	Max	$\bar{X}/\text{Max}$
200	135	321			268	365	357	312	232	7	284	400	.71
150	110	330	283	207	195	302	290	227	288	9	248	"	.62
100	45	202	244	192	215	347	242	181	226	9	210	"	.53
50	19	171	108	49	50	278	93	52	164	9	109	"	.27
30	13	90	32	29	25	72	28	45	110	9	49	300	.16

the train, and hence the total number of presses in a fixed interval of time might be greatly influenced by this purely arithmetical fact; for example, the 30/second stimulus train has a duration of 1 second. In view of this fact, it became desirable to construct an index that would avoid the possible confounding and reflect more validly the rat's incentive to press when not being stimulated. For this purpose we decided upon a maximum pressing rate of 300 for a 5-minute period when the stimulus repetition rate was 30 pps (if a rat pressed immediately after each 30/second stimulus train and at no other time, he would press 300 times in 5 minutes), and of 400 for all other repetition rates (365 presses was the highest number achieved by any rat under any condition). By forming the ratio "average number of actual presses/maximum number," we obtained an "activity ratio" for each repetition rate. This is given in the " $\bar{X}/\text{Max}$ " column. It will be noted that pressing activity falls sharply below 100 pps.

It is interesting to do the analogous experiment with duration of stimulus train constant (.2 second) and number of pulses variable. This was performed on only three of the animals and the results are shown in Table 2. It will be noted here that the fall in pressing rate below 100 pps is even sharper.

In the above experiments, no attempt was made at a factorial design; each rat was tested in the order 200, 150, 100, 50, 30 pps, with all repetition rates presented in one session, but with a 5-minute rest between repetition rates. At the beginning of each session the rat was given "prompting" stimulations if necessary and allowed to press for 60 seconds before records were taken.

c. *Rate of lever pressing as related to pattern of evoked rhythmic after-*

TABLE 2

NUMBER OF PRESSES IN 5-MINUTE PERIOD AS RELATED TO STIMULUS REPETITION RATE: NUMBER OF PULSES PER STIMULUS TRAIN, VARIABLE; TRAIN DURATION, CONSTANT (.2 SECOND)

PPS	# 759	# 762	# 763	N	$\bar{X}$	Max	$\bar{X}/\text{Max}$
200	284	204	238	3	242	400	.61
150	337	220	227	3	261	"	.65
100	226	212	178	3	205	"	.51
50	42	45	27	3	38	"	.10
30	17	28	6	3	17	"	.04

*potential.* Of paramount interest is the covariation between the energetics of lever pressing and the average evoked potentials generated in the process. To this end, a fixed interval of about 2.25 seconds between triggered sweeps was utilized, so that the 2-second sweep on the averaging apparatus plus the duration of the stimulus train would intervene between successive recorded presses. In this experiment we utilized the fixed train duration of .2 second for the various repetition rates. The dependent variable was the amount of time required of the rat to generate 128 sweeps. The chief drawback in this maneuver is the fact that with repetition rates below 100 pps, the animals very often pressed in a desultory fashion or not at all, forcing *E* to administer prompts. Thus the time values achieved for these rates become meaningless. In general, *E* waited about 10-20 seconds before giving a prompting stimulation. The minimal time required to generate 128 sweeps would be 288 seconds ( $128 \times 2.25$ ); any score in excess of about 800 seconds involves varying numbers of hand stimulations and is merely indicative of a very low pressing rate.

The tracings, representing 128 sweeps, and the times required to generate them are given in Figure 4 for the three animals used. It is seen that a nice correspondence exists between the amplitude and duration of the rhythmic



potentials and liveliness of lever pressing. When the experiment was performed under the condition of constant pulse number (30 pulses) and variable stimulus train durations, the same result is obtained, but the drop-off in oscillations for repetition rates of 50 and 30 is not so sharp.

*d. Covariation between presence-absence of lever pressing and presence-absence of oscillations.* The preceding data indicate that (a) the characteristics of the oscillations when they exist are related to the repetition rate of the stimulus; (b) when the rat is disposed to press for intracranial stimulation, the rate of pressing is also related to the repetition rate of the stimulus; and (c) when disposition to press and repetitive potentials both are present, the rate of pressing and the parameters of the waves vary together. It is desirable to know whether, in the context of the present investigation, the neurophysiological process or the behavioral process can exist without the other. In other words, is it ever the case that a rat shows oscillatory afterpotentials to stimulation even though he refuses to press the lever; or that a rat will press for stimulation without showing oscillatory potentials?

Altogether 22 rats were used in the present investigation; of these, 11 became good pressers, 11 did not. On the basis of visual inspection of traces averaging not less than 128 sweeps obtained under "optimal" stimulating conditions—i.e., repetition rates of 200 or 150 pps and under the 30 pulses per train criterion—these 22 traces were divided into two groups: (a) those showing oscillatory potentials and (b) those not showing. This judgment was quite easy to make, as can be seen by inspecting the tracings in Figures 3, 4, and 5.

The material shown in Table 3 is the result.<sup>2</sup> The covariation is quite good; there are two exceptions. These animals gave undulatory potentials—not strong or numerous—but never learned to press, even after many hours spent in shaping. Stimulation (by hand) caused them to "freeze," and if repeated, they settled down into what appeared like a drowsy or sleeping state—although when returned to the home cage they became immediately lively. When the forepaws were placed on the lever, they were immediately removed, and in what little wandering occurred, the animals kept away from the lever. They never sniffed or "showed interest" in the surroundings and were much like the other nonpressing rats, except that stimulation at 200 or 150 pps caused 2-3 undulations of small amplitude but of average periodicity.

Animals who gave good repetitive undulations from a given pair of pick-up electrodes if stimulated by 1-2, showed flat tracings when stimulated by other

<sup>2</sup> Identical results were obtained when a two-way classification was made on the basis of tracings generated by hand stimulation and in the absence of the lever.



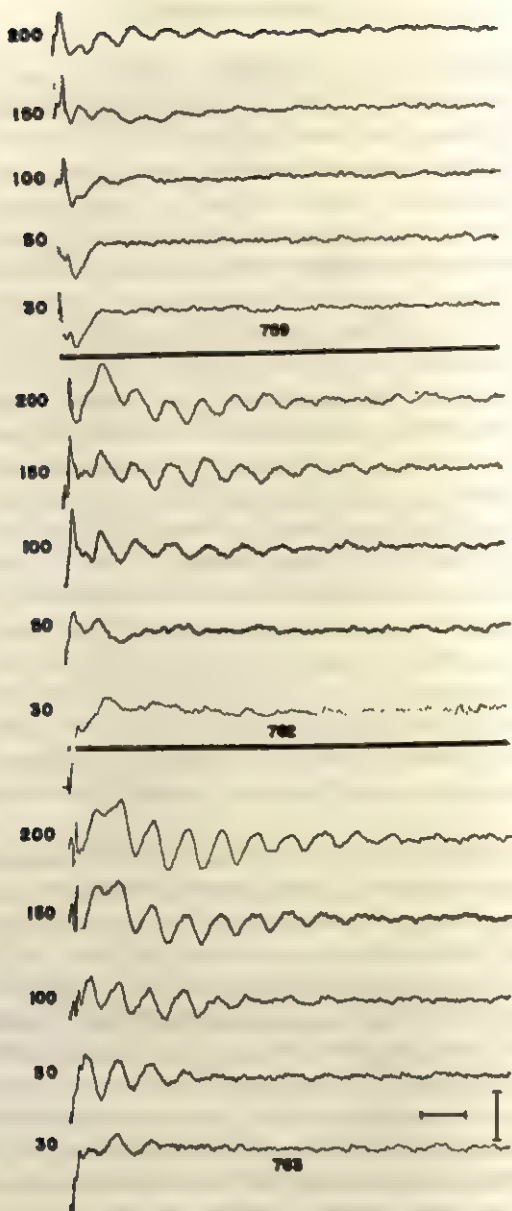


FIGURE 4

TRACING FROM THREE ANIMALS, SHOWING RELATION BETWEEN STIMULUS REPETITION RATE (TRAIN DURATION FOR ALL RATES CONSTANT AT .2 SECONDS) AND THE TOTAL TIME REQUIRED FOR THE S TO GENERATE 128 SWEEPS FOR EACH REPETITION RATE

Identity of rat indicated at bottom of each set; the repetition rates, in the left margin. The corresponding time, in seconds, for each trace is as follows: #759, 401, 422, 556, 810, 869; #762, 390, 404, 480, 756, 1080; #763, 390, 481, 453, 629, 1140.

Each trace is 2" in duration. Stimulation amperage kept constant for each rat at all repetition rates, but varied from one rat to another.

Calibrations: in this and all other  $\bar{X}$ - $\bar{Y}$  plots, vertical line = 50  $\mu$ V; horizontal line = 200 msec.

TABLE 3  
COVARIATION BETWEEN AFTER-DISCHARGES AND SELF-STIMULATION  
Presence of periodic after-discharges

		Yes	No
Pressing for	Yes	11	0
intracranial			
self-stimulation	No	2	9

N = 22

Note:  $\chi^2 = 16.2$ ;  $p < .01$ .

electrode pairs. However, in these latter cases the stimulating electrodes were separated by 2.5 mm or more and involved large segments of the cerebrum; hence voltage had to be decreased drastically in order to avoid massive somatic reaction. In view of this fact, the resulting tracings cannot be evaluated.

In this connection, one experiment was carried out in which each electrode assembly was provided with two pairs of stimulating electrodes; i.e., the two electrodes in each pair were parallel and juxtaposed and of approximately the same length. These were the pairs 1-2 and 5-6, with 5-6 being given the same configuration as 1-2. Two animals were used. In one S (#762), the 5-6 electrodes were 6.0 and 6.5 mm in length; in the other (#763), electrodes 5-6 were 9.0 and 9.5 mm in length. It was considered that the 5-6 electrodes in #762 were too short to be effective, and those in #763 were too long. The 1-2 electrodes were of the usual length for both animals.

The averaged traces for #763 are shown in Figure 5. It will be seen that the rhythmic, repetitive potentials are very pronounced when stimulation is by 1-2, whether the pick-up be from 3-4 or 7-8. That is,  $St_{1,2}$   $P_{3,4}$  and  $St_{1,2}$   $P_{7,8}$  both yield rhythmic patterns. The animal's concomitant behavior was one of frantic pressing. When he was stimulated through 5-6, on the other hand, the repetitive discharges were missing. Accompanying this latter electrographic fact was a disinterest in the lever under  $St_{5,6}$ , such that the stimuli had to be delivered by E.

For #762, however, there were vigorous undulations of potential and very

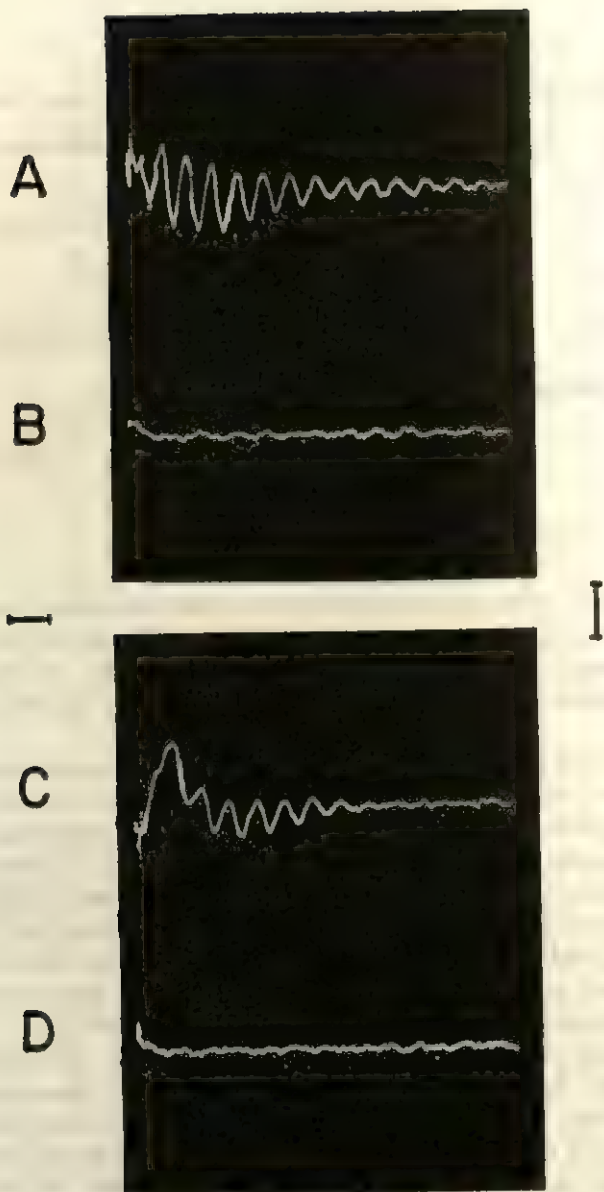


FIGURE 5

TRACINGS OBTAINED FROM THE SAME RAT: STIMULUS VOLTAGE CONSTANT;  
REPETITION RATE CONSTANT AT 200 PPS

(A) Stimulating electrodes are the usual  $E_1-E_2$ , the pick-up electrodes are  $E_3-E_4$  [ $St_{1-2} P_{3-4}$ ]; (B) is  $St_{3-6} P_{3-4}$ ; (C) is  $St_{1-2} P_{7-8}$ ; and (D)  $St_{5-6} P_{7-8}$ . In this experiment  $E_5-E_6$  are bent to form the same configuration as  $E_1-E_2$ .

Stimulation through the 1-2 electrodes (aimed at the medial forebrain bundle) is effective in producing repetitive after-discharges, while stimulation through 5-6 is not. This is true both for  $P_{3-4}$  and  $P_{7-8}$ .

All four tracings were generated by hand stimulation and based on 256 sweeps.

eager pressing of the lever for self-stimulation of both  $St_{1,2}$  and  $St_{5,6}$ ; apparently, electrodes 5-6 were not short enough.

*e. Homolateral vs. heterolateral evoked potentials.* We shall concern ourselves here exclusively with potentials evoked by stimulating electrodes 1-2. Potentials recorded when both pick-up electrodes are on the same side of the midline will be termed homolateral; those recorded when the two pick-up electrodes are on opposite sides, heterolateral. The former are 3-4, 5-6, 5-7, 5-8, 6-7, 6-8, 7-8; the latter are 3-5, 3-6, 3-7, 3-8, 4-5, 4-6, 4-7, 4-8. Since no histological verification of electrode placement was made, comparison of the tracings coming from these two classes of electrode placements might be of some interest. Exhaustive and systematic observations were not made: only a sampling of the two classes of electrode placement. No obvious differences in amplitudes were evident to mere visual inspection. Counts were made of the number of consecutive waves for each of 62 records; 35 of these records involved heterolateral pairs of electrodes, and 25 homolateral. Here, again, there were no obvious differences; the average number of consecutive waves for the heterolateral was 4.2, and for the homolateral 5.0. The judgment as to what constituted a "wave" was purely subjective.

*f. Tracings obtained from self-stimulation vs. those from priming stimulations.* Throughout the investigation, averaged tracings were obtained under both conditions; under the latter, the lever was not in the behavior compartment. No consistent differences were observed except perhaps in the first 200 msec after the beginning of the averaged trace. During this interval the first wave sometimes showed an exaggerated amplitude and more irregular shape for some animals and some electrode pairs.

*g. Averaged brain potentials during the course of the stimulus train.* On three of the animals used in this investigation, an experiment was performed in which the oscilloscope sweep was triggered at the beginning of the stimulus train. The biasing device already described made it possible to keep stimulus artifact gratifyingly small. In these experiments the duration of the stimulus train was kept at 600 msec for all repetition rates used. The train was made quite long in order to get a better "view" of evoked potentials during the period of stimulation. Only two repetition rates were employed: 150 and 50 pps, with other stimulus parameters fixed in the usual manner. The stimulating voltage, of course, was constant for a given animal.

The results are of special interest in that the stimulus train at 150 pps did not evoke well-defined waves during its course, whereas waves were somewhat more in evidence at 50 pps, although clearly so in one case only. The averaged tracings for the three animals are shown in Figure 6. It will be

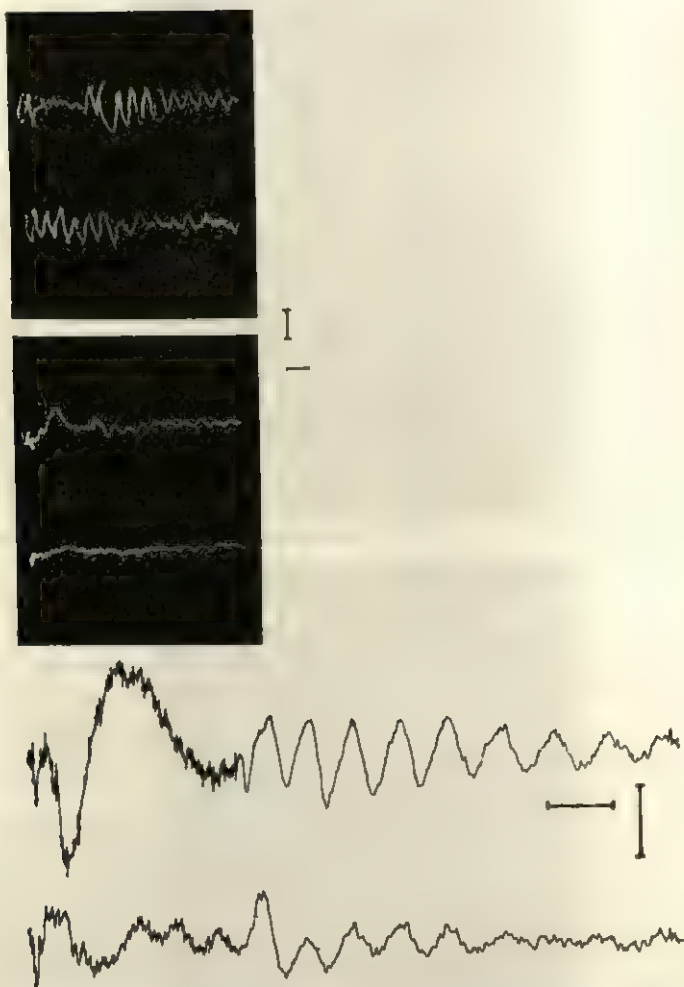


FIGURE 6

THREE PAIRS OF TRACES, EACH PAIR BELONGING TO A DIFFERENT RAT

The upper trace of each pair is generated by pulses at 150/second, the lower trace by pulses at 50/second. The duration of the stimulus train is .6" in all cases, and the sweep is initiated at the *beginning* of the stimulus train. The first part of the sweep, therefore, records averaged brain potentials during stimulation. Hand stimulation.



seen that whereas the poststimulus segments show the characteristic repetitive after-discharges for 150 pps, the evoked potentials during the stimulus train vary considerably. The tracings shown in Figure 6 were all taken with  $St_{1,2}$   $P_{7,8}$ . Analogous tracings were obtained with  $St_{1,2}$   $P_{3,4}$ ; the patterns of evoked potential during stimulation were generally similar to those issuing from  $P_{7,8}$ , in that waves were marginal or absent.

This absence of patterned potentials during the application of an "optimal" stimulus repetition rate is rather surprising until it is considered that at 150 pps an "empty" time interval of about 6.2 msec intervenes between successive pulses. Thus, the very rapidity of stimulation interferes with the development of the reverberation phenomenon, since the period of the undulations is 120-160 msec. At the repetition rate of 50 pps matters are somewhat but not much better.

*h. Incidental electrographic observations.* Additional observations were noted, as follows:

(a) Of the 13 animals that gave undulation of potentials under optimal conditions to one pair of pick-up electrodes, only four failed to give undulations to at least one other pair. This uniqueness of the four in question may only reflect the fact that an exhaustive study of each possible pair of pick-up electrodes was not made for each of the animals used. It was found that if no waves were forthcoming on either  $P_{3,4}$  or  $P_{3,8}$  or  $P_{7,8}$ , none of the other pairs would prove fruitful.

(b) While no systematic study was made of this point, it was frequently observed on the monitoring oscilloscope that an animal that had had extensive positive reinforcing stimulation gave an EEG that showed a marked predominance of high amplitude, low frequency waves, even though no stimulus was being administered and no lever was in the apparatus. By contrast, when the animal had finished a session of nonreinforcing stimulation, the low frequency, high amplitude waves were not so prominent. In one case, at least, they seemed completely absent after a 30 pps session and very pronounced after a 200 pps session on the same day.

(c) Related observations in the above context are the following: After a period of 5-6 idle days, some practiced, good presser animals showed only a mild interest at the beginning of an experimental session; they would make no attempt to press the lever and would wander around the apparatus in a casual fashion. When hand-controlled stimulation was administered, there was only a mild sniffing reaction. This state of affairs might last over a period of 8-10 stimulations. During this short period the averaged traces would show very little rhythmic undulations. As time went on, with repeated hand

stimulations, these animals would become excited and begin eagerly to seek and press the lever for self-stimulation. The averaged tracings for a comparable number of sweeps now would show definite evidence of potential oscillations. This phenomenon was observed whether the lever was present or absent in the container.

(d) If the rat is placed under deep ether anesthesia (disappearance of tail reflex, etc.), the rhythmic phenomena present under optimal stimulating conditions in the awake animal disappear. In these experiments I was careful to place the animal in a prone position, such that the soles of his feet made contact with the floor, thus simulating somewhat the electrical situation obtaining normally. A stimulus repetition rate of 200 or 150 pps, giving an averaged trace with beautiful, rhythmic waves in an awake animal, characteristically gave a flat trace in the etherized state. In one case, stimulation at 200 pps normally gave six waves during a time interval of 1 second; the same animal, etherized, but with identical stimulus parameters, gave but two waves of diminished amplitude immediately upon termination of the stimulus train; the rest of the trace (1555 msec) was completely flat.

#### D. DISCUSSION

Perhaps the most noteworthy thing about the obtained oscillations is their regularity, symmetry, and form; with 256 sweeps (Figure 3), the smoothness of the diminishing periodic pattern is suggestive of a condenser discharging through an inductance and is somewhat alarming to an experimenter. However, in the several experiments with anesthetized animals (ether), the characteristic waves did not appear at any repetition rate of stimulation. In these experiments great care was taken to place the animal in a prone position in such a way that the plantar surfaces of his feet were in contact with the bottom of the container. Furthermore, the disappearance under normal conditions of the undulations with diminishing repetition rates of the stimulus, and their failure to appear when stimulating electrodes are not placed in effective areas, does not give support to the hypothesis of a purely physical artifact.

The great constancy of the wave period, 120-160 msec (8-6 cycles/second), in different animals and in the same animal in different segments of the tracings is also impressive, as, of course, is the fact that this range of periods is in the zone of alpha frequency. Worthy of note, too, is the fact that periodic after-potentials could be observed in several regions of the cerebral mass in response to a given stimulus train applied to a given locus. While exhaustive explorations with all six recording electrodes were not undertaken,

enough data are at hand to show that the oscillatory phenomenon is fairly widespread in the rat's brain. Thus, it was very common to record periodic discharges from  $P_{3,4}$ ;  $P_{3,7}$ ;  $P_{3,8}$ ;  $P_{4,8}$ ;  $P_{7,8}$ , when the repetition rate (200 or 150 pps) and the electrode placement were effective, although the number of waves and their amplitude were not identical for the various sources.

The phenomenon of periodic potentials in the alpha range has been noted a number of times by various investigators working with average evoked responses to light in man [for references to literature, see Lansing and Barlow (8)]. Lansing and Barlow (8) have investigated the relationship between the time course and amplitude of the periodic after-discharges and the parameters of the ongoing alpha rhythm at the moment of visual stimulation. The tracings given by them of the after-discharges associated with visual evoked potentials and picked up by scalp electrodes have the same regularity and approximate period as those reported in the present paper, and the train of diminishing waves is of the same duration, of the order of 1-2 seconds. A curious detail noted in some of our tracings also appears in their paper; viz., the virtual disappearance of the undulation after about 1 second for a period of 200-400 msec, and then its reappearance until the end of the sweep with diminished amplitude (see their Figure B, D).

The periodic after-discharge phenomenon—sometimes referred to as “ringing” in laboratory jargon—has also been observed when stimulation involves other than the visual modality, and the pick-up is made directly from nerve tissue in acute preparations. Adrian (1) discovered in 1941 that a single tactile stimulus elicited a series of waves in the thalamus of about 10/second which he called “thalamic after-discharges.” Such peripheral stimuli elicited a series of 3-7 such cycles, and by recording simultaneously from the white matter below the cortex, he showed that a similar rhythmic discharge occurred in the thalamo-cortical fibers. Andersen *et al.* (3) generated a train of rhythmic discharges in the thalamus, 8/second, with a single afferent volley to the contralateral radial nerve, with similar discharges being observed simultaneously from the killed ends of the thalamo-cortical fibers. The repetitive nature of the thalamic response persisted after the removal of the appropriate cortical projection area.

The experiments reported by us differ from those, for example, of Adrian and Andersen in that we are dealing with chronic, free-moving animals, the stimulation is to the medial forebrain bundle (presumably) or adjacent structures, and the undulations are the result of an averaging process. However, the basic characteristics of our waves—namely, smoothness, period, duration, and diminishing amplitude—impel one to conjecture that we are dealing here

with a basic membrane rhythmicity that manifests itself upon adequate stimulation (2).

Our primary concern as psychologists is whether the undulations of potential, when generated by appropriate stimulation of certain critical areas, are intrinsic features of the neurophysiology of motivation; that is, whether they constitute the necessary, if not sufficient, condition for positive reinforcement. The data obtained within the framework of the present experiments seem to point in this direction.

The experimental results of Hess *et al.* (7), Clemente *et al.* (4), and Roth *et al.* (10) are very suggestive in this regard. Hess had noted bursts of 5-8 cycles/second, high voltage EEG waves in cats while they were drinking milk. Clemente recorded, in an experimental conditioning situation, bursts of EEG waves, 100-150  $\mu$ V and 4-8 cycles/second, after cats had pressed a manipulandum and were in the process of consuming the reward (mixture of milk and broth). As the animals became overtrained, the bursts became more definite and less diffuse. After the reward was consumed, the EEG immediately showed low voltage, fast activity. The synchronization bursts appeared simultaneously in leads (screws) placed permanently in the calvarium over widespread cortical areas. Clemente calls the phenomenon "post reinforcement synchronization." On substitutions of water for the customary mixture of milk and broth, the post reinforcement synchronization was abolished and replaced by the usual low amplitude, high frequency activity. Concurrently, the animals, after sampling the "reward" several times, would pace the floor and return from time to time to press the manipulandum. Similarly, Roth obtained from cats bursts of slow frequency (4-12/second), high amplitude waves during and briefly subsequent to consumption of milk reward for pressing a lever. The potentials were derived from screw electrodes placed subdurally.

The range of frequencies observed by these investigators went lower than those recorded in the present report. Whether this is attributable to the different species of animal used, the method of deriving the waves (single records *vs.* averages), the locus of the pick-up electrodes, or the nature of the applied stimulus, is undetermined. The obvious similarities of all these observations are, however, strongly suggestive of a common relation to the dynamics of reinforcement.

The reason for the marked covariation in the present investigation between the degree of rhythmicity in the after-potentials and the repetition rate of the applied stimulus is not clear. Is the correlation merely an expression of the phenomenon of temporal summation? That is, is it the case that



the repetition rate of 200 pps, for example, evokes a more pronounced undulation because any stimulus in the train after the first is superimposed on the residual depolarizations caused by the preceding ones and is therefore a more intense stimulus, whereas in the repetition rate of 30 pps the residual depolarization largely disappears in the interval between pulses?

Temporal summation is undoubtedly a factor in this situation but probably not the dominant one. In Figure 7A are shown the averaged traces given by one animal with 200 pps and 5 V; in 7B the same animal, 1 pps and 15 V. (There was no visible somatic reaction at the later voltage.) In both conditions the stimulus train was composed of 30 pulses; the interval between successive stimulus trains was randomly variable but averaged 10 seconds. It will be seen that A has a definite, if somewhat short, train of rhythmic waves, whereas B lacks rhythmicity. The concomitant behavior of the animal in A was one of mild interest, and only 30 percent of the sweeps were generated by the animal's pressing the lever; in B the animal became increasingly quiet, and although the lever was present, practically all the stimulations were delivered by *E*.

Although in this case it is impossible to ascertain objectively, it is not easy to believe that the factor of temporal summation would completely transcend the effect of a 300 percent difference in voltage between (A) and (B).

The averaged tracing 7C is from the same animal when stimulated by a single pulse at 15 V with no lever present. The interval between successive stimulations varied between 15 and 20 seconds. The evoked potential here is highly irregular although definite enough. In view of the strength of the stimulus applied, one would expect some evidence of the rhythmic phenomenon without benefit of temporal summation.<sup>3</sup> Line C should be compared with D, which is a control tracing where no stimulus was applied. During C the animal was very quiet and remained in one position, while in D there was some casual walking around. In trace E the animal stimulated himself at 200 pps (30 pulses per train) and 8 V. The difference between it and the other four tracings is obvious. Not only are there more waves in E than in A, they are also of greater amplitude. However, it will be noted that their period is the same. Moreover, the behavior of the same animal during A and E was spectacularly different; in A there was only mild interest and desultory pressing, in E the lever was attacked in a very energetic and al-

<sup>3</sup> While the factor of temporal summation is probably of secondary importance in initiating impulses in those fibres that are adjacent to the stimulating electrodes, it could be critically involved in neurones "downstream" that are members of reverberating circuits, since for these latter elements many synaptic junctions are involved.



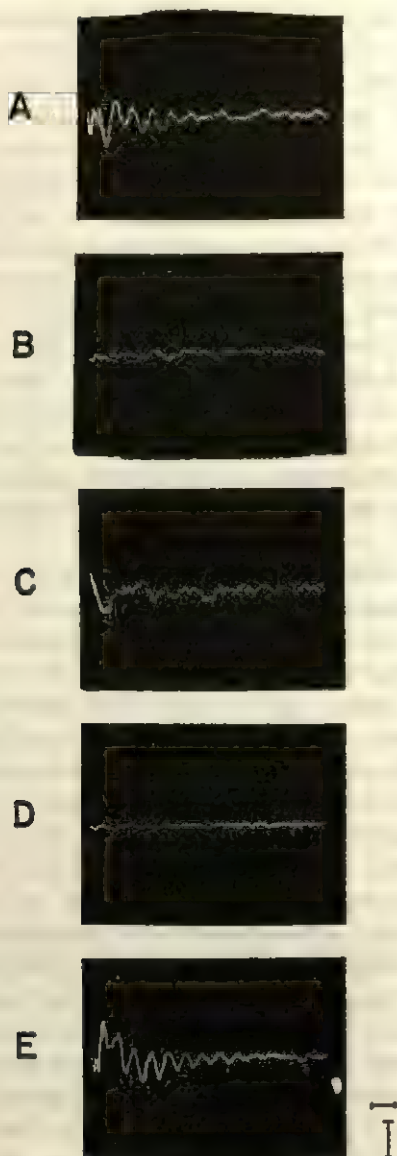


FIGURE 7

ALL TRACINGS FROM THE SAME RAT

(A) 200 pps, 5 V, stimulus train 30 pulses; (B) 1 pps, 15 V, stimulus train 30 pulses; (C) sweep generated by a single pulse, 15 V (i.e., the "train" was constituted of a single pulse applied at irregular intervals of 15-20 seconds); (D) control (no stimulation); (E) 200 pps, 8 V, stimulus train 30 pulses, self-stimulation.

most frantic manner. Two other animals, similarly tested, gave identical results.

It is worthy of comment that we did not obtain the "ringing" phenomenon reported by Lansing and others (1, 3, 8, 11) in response to the application of a single stimulus volley. In two of the three rats used there was some initial evidence of rhythmicity at the beginning of the second half of the 2-second sweep, but this was washed out in the average of 128 sweeps.

A consideration that arises from the foregoing data involves the fact that the observed periodicities have been picked up with bipolar recording. Since both pick-up electrodes were imbedded in the cerebral mass, a certain ambiguity occurs in formulating an image of the space-time distribution of underlying neural events. It is desirable to secure monopolar recordings under identical stimulating conditions, where one of the recording electrodes is placed at a "neutral" point.

An even more important consideration relates to sites in the cerebral mass whose stimulation is aversive. In the experiments reported here, we can only say that the tracings were flat—with the marginal exception of two animals—in rats that did not press for stimulation at any stimulus repetition rate. We conclude, therefore, at least tentatively, that cerebral self-stimulation is accompanied by characteristic electrographic signs. The question that suggests itself is this: What are the electrographic signs accompanying a stimulus that *inhibits* ongoing cerebral self-stimulation? Flatness of the averaged tracing cannot be such a sign, since this merely means that no systematic neural event was generated by the stimulus.

In summary, it is hard to escape the thought that we are dealing here with a membrane resonance phenomenon, where the period of resonance is governed by the membrane's molecular structure and hence is maximally responsive to a limited range of stimulus repetition rates (9). Moreover, since rats (and other animals) will activate any manipulandum (rotate a wheel, press a lever, choose a particular route, etc.) to give themselves suitable cerebral stimulation, we must suppose that appropriate stimulus repetition rates applied at the "correct" locus produce a rhythmic neural process sufficiently widespread in the cerebral mass as to be capable of influencing plastic changes in any one of many neural circuits. Should this hypothesis be true, however (5, 9), the relationship between these periodic after-effects and the physiology of neurotransmitters and axon spikes as they relate to behavior change would still remain to be elucidated.

#### REFERENCES

1. ADRIAN, E. D., Afferent discharges to the cerebral cortex from peripheral sense organs. *J. Physiol.*, 1941, **100**, 159-191.

2. ANDERSEN, P., & ANDERSSON, S. A. *Physiological Basis of the Alpha Rhythm*. New York: Appleton-Century-Crofts, 1968.
3. ANDERSEN, P., BROOKS, C., ECCLES, J. C., & SEARS, T. A. The ventro-basal nucleus of the thalamus: Potential fields, synaptic transmission and excitability of both pre-synaptic and postsynaptic components. *J. Physiol.*, 1964, **174**, 348-369.
4. CLEMENTE, C. D., STERMAN, M. B., & WYRWICKA, W. Post-reinforcement EEG synchronization during alimentary behavior. *Electroenceph. & Clin. Neurophysiol.*, 1964, **16**, 355-365.
5. ELUL, R. Brain waves: Intracellular recording and statistical analysis help clarify their physiological significance. In *Data Acquisition and Processing in Biology and Medicine* (Vol. 5). New York: Pergamon Press, 1968.
6. GENCERELLI, J. A., PRIDDY, R. D., & AVERILL, J. R. Studies in the neurophysiology of learning: VII. Bar-pressing rate for cranial self-stimulation as related to frequency of rectangular pulses when  $Q$  is held constant. *J. of Psychol.*, 1963, **56**, 455-472.
7. HESS, R., JR., KOELLA, W. P., & AKERT, K. Cortical and subcortical recordings in natural and artificial sleep in cats. *Electroenceph. & Clin. Neurophysiol.*, 1953, **5**, 75-90.
8. LANSING, R. W., & BARLOW, J. S. Rhythmic after-activity to flashes in relation to the background alpha which precedes and follows the photic stimuli. *Electroenceph. & Clin. Neurophysiol.*, 1972, **32**, 149-160.
9. MONNIER, A., & COPPÉE, G. Relation entre la rhythmicité de la réponse nerveuse et la résonance. *Arch. Internat. Physiol.*, 1939, **48**, 129-180.
10. ROTH, S. R., STERNMAN, M. B., & CLEMENTE, C. D., Comparison of EEG correlates of reinforcement, internal inhibition and sleep. *Electroenceph. & Clin. Neurophysiol.*, 1967, **23**, 509-520.
11. TEIL, S., FELBER, F., & KLINGBERG, F. Der Einfluss der mesenzephalon retikulären Formation auf die Rhythmenbildung im Corpus geniculatum der Ratte. *Acta Biol. & Med. German.*, 1972, **28**, 637-649.

*Department of Psychology*

*University of California, Los Angeles*

*Los Angeles, California 90024*

## LOCUS OF CONTROL AND SELF-DISCLOSURE OF PUBLIC AND PRIVATE INFORMATION BY COLLEGE MEN AND WOMEN: A BRIEF NOTE\*

*Department of Psychology, University of Maine at Orono*

RICHARD M. RYCKMAN, MARTIN F. SHERMAN, AND GARY D. BURGESS

Jourard<sup>1</sup> has argued that American college students who disclose little information about themselves to others tend to be maladapted. Although internally oriented people<sup>2</sup> tend to behave in irrational and self-defeating ways under some circumstances,<sup>3</sup> the bulk of the literature concerned with individual differences in feelings of powerlessness indicates clearly that externally oriented people have a generally poorer conception of their personal worth,<sup>4</sup> are more anxious,<sup>5</sup> and behave more inappropriately and unrealistically on tasks that demand the use of their skills.<sup>6</sup> Consequently, a study was conducted in which it was predicted that externals would report that they disclose less about themselves to their parents and to male and female friends than internals, regardless of the intimacy level of the information being revealed. In addition, an attempt was made to replicate the findings of previous studies<sup>7</sup> which indicated that women tend to disclose more information about themselves than men and that study participants would disclose the least amount of information to their fathers than to the other target groups.<sup>8</sup>

To test these predictions, 80 students at the University of Maine and the California State University at Los Angeles completed Rotter's Internal-

\* Received in the Editorial Office on May 7, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Jourard, S. M. *The Transparent Self* (2nd ed.). New York: Van Nostrand, 1971.

<sup>2</sup> For a complete description of the locus of control construct, see Rotter, J. B. Generalized expectancies for internal *versus* external control of reinforcement. *Psychol. Monog.*, 1966, **80**, Whole No. 609.

<sup>3</sup> Ryckman, R. M., & Sherman, M. F. The effects of success and failure and locus of control on the selection of a partner or opponent in a problem-solving situation. Paper presented at the Eastern Psychological Association, Washington, D. C., 1973.

<sup>4</sup> Fish, B., & Karabenick, S. A. Relationship between self-esteem and locus of control. *Psychol. Rep.*, 1971, **29**, 784.

<sup>5</sup> Butterfield, E. C. Locus of control, test anxiety, reaction to frustration, and achievement attitudes. *J. Personal.*, 1964, **32**, 298-311.

<sup>6</sup> See Lefcourt, H. M. Internal *versus* external control of reinforcement revisited: Recent developments. Research Report No. 27, University of Waterloo, Waterloo, Ontario, Canada, 1971.

<sup>7</sup> Cozby, P. C. Self-disclosure: A literature review. *Psychol. Bull.*, 1973, **79**, 73-91.

<sup>8</sup> See footnote 1.

External Scale<sup>9</sup> and a modified version of Jourard's Self-Disclosure Scale<sup>10</sup> in their classrooms and dormitories. The revised questionnaire contained 24 items, 12 of which were items of a public or nonintimate nature, covering disclosure of information about attitudes and opinions, tastes and interests, and academic studies, and 12 items of a private or more intimate nature.<sup>11</sup> These latter items included information about money, personal adequacy, and feelings about one's body.

The results indicated that women generally disclosed more self-information than men and that both college men and women reported disclosing the least amount of information to their fathers. Also, as expected, externals ( $M = 93.04$ ) had a tendency to disclose less information about themselves to intimates than internals ( $M = 107.52$ ), regardless of the public or private nature of the information ( $F = 3.70$ ,  $df = 1, 76$ ,  $p < .10$ ). On the basis of a maximum self-disclosure score of 192, externals disclosed 48% of all possible information about themselves, while internals disclosed 56%. Caution should be exercised in interpreting these results, however, as indicating that externals are more maladapted than internals, since Jourard's argument about the nature of the relationship between self-disclosure level and maladaptiveness hinges upon the meaning of the term "little" disclosure. It seems unlikely that the 48% disclosure of externals to intimates would reflect such a minimal level, but a more precise answer to this problem awaits the outcome of future research.

*Department of Psychology*  
*University of Maine at Orono*  
*Orono, Maine 04473*

<sup>9</sup> See footnote 2.

<sup>10</sup> See footnote 1.

<sup>11</sup> See Fitzgerald, M. P. Self-disclosure and expressed self-esteem, social distance and areas of the self revealed. *J. of Psychol.*, 1963, **56**, 405-412.



## MMPI TWO-POINT CODES FOR A "NORMAL" COLLEGE MALE POPULATION: A REPLICATION STUDY\*<sup>1</sup>

Ohio University

---

IVAN GULAS<sup>2</sup>

---

### SUMMARY

Two-point MMPI code distributions for a sample of 609 precollege students were compared with the code distributions of 701 precollege students obtained two years previously. The samples showed highly similar two-point code distributions. The same seven two-point code types occurred most frequently in both samples, and these same seven code types accounted for over 50% of the cases in both samples.

Further support was thus obtained for the hypothesis that Scale 5 is related more to *IQ* and/or education aspiration and/or socioeconomic status than years of education *per se*. Also, the original base rate data on two-point code frequencies, including Scales 5 and 0, were replicated, providing reliable data on "normal" college males. Previous research of this nature with this population has typically omitted Scales 5 and 0.

---

### A. INTRODUCTION

Configural interpretation of the Minnesota Multiphasic Personality Inventory (MMPI) has significantly contributed to the utility of this instrument, facilitating greater interpretive precision. Configural research on base rates of MMPI two-point codes has revealed, somewhat surprisingly, that relatively few two-point code types account for the majority of cases sampled. Webb (6) reported that on a nationwide sample of 12,174 psychiatric patients, eight two-point code types (including Scales 5 and 0) accounted for over 50% of the male cases. Hathaway and Meehl (5) and Guthrie (4) reported that 60% of the cases in their study fell into nine of the 56 digit code classes (with Scales 5 and 0 omitted). The omission of these two so-called "nonclinical"

---

\* Received in the Editorial Office on May 7, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The author is indebted to James T. Webb, Ph.D., Ohio University, for his assistance in preparation of this manuscript, and to F. W. King, Ph.D., Dartmouth College, for making the use of the Dartmouth MMPI data available.

<sup>2</sup> Requests for reprints should be sent to the author at the address shown at the end of this article.

scales by these investigators is unfortunate in light of recent research developments. Gulas (3) found that seven two-point code types accounted for over 50% of the cases in a precollege sample, and each of these seven two-point codes included an elevation on Scale 5.

The relationship between two-point code types and demographic variables has been investigated by Aaronson (1), Gough (2), Webb (6), and Gulas (3). Substantial evidence now exists that age, sex, and education strongly influence the distribution of high-point MMPI codes, and that these parameters should be considered as moderator variables in clinical interpretation of the MMPI.

The relationship between elevation on Scale 5 and education has been reported by several investigators (1, 2, 6). Gulas (3) hypothesized that Scale 5 is related more to *IQ* and/or educational aspiration and/or socioeconomic status than to years of education *per se*. This hypothesis was further evaluated in the present study.

Gulas (3) compared his normal precollege sample with Webb's (6) psychiatric sample on the basis of the two-point code frequency distributions. Potential clinical interpretations were hypothesized where disproportionate frequencies of two-point codes existed between the two populations. The validity of those interpretations is, in part, contingent on whether the college sample two-point code distribution reported by Gulas (3) represented an unbiased sample of the male college population. The present study provided a replication and extension by comparing the previously reported two-point code distribution for college males (3) with an independent college male sample, thus providing a check on possible sampling errors in the original study.

## B. PROCEDURE

Data gathering of the present study followed the identical procedure previously utilized by Gulas (3). Computer processed MMPI profiles on 609 males (one year's entire graduating class of Dartmouth College) were obtained and constituted the Dartmouth II sample. As in the Dartmouth I sample (3), the *Ss* took the MMPI during their freshman orientation week, typically four years prior to their graduation. Each of the 609 MMPI profiles was converted to two-point codes.<sup>3</sup> Since each *S* was successful in graduating, it is assumed that this represents an intellectually well-functioning sample of young adults. Comparisons of the two-point code frequencies between this Dartmouth II sample and the Dartmouth I sample of 701 male precollege

<sup>3</sup> Two-point codes were defined as the highest and the second highest of the MMPI clinical scales regardless of the absolute *T*-score value of the scales.

students (obtained two years prior to the Dartmouth I sample) were computed by means of chi-square analyses.

### C. RESULTS AND DISCUSSION

#### 1. Normative Data

The two-point codes were arranged in decreasing order of frequency of occurrence. The results are presented on Table 1. Inspection of Table 1 reveals that seven two-point code types account for 51.5% of the cases in this Dartmouth II sample. Interestingly, these same seven code types had accounted for 51.6% of the Dartmouth I sample previously reported by Gulas (3). Thus codes 35/53, 59/95, 25/52, 57/75, 58/85, and 56/65 appear to be consistently the seven most frequent code types for college males, and account for over 50% of the code types.

Since the seven most frequent code types in the two samples are identical, and since each two-point code includes an elevation on Scale 5, the significant relationship of Scale 5 to demographic variables is strongly substantiated. Of these seven two-point codes, Webb (6) found six of them (with the excep-

TABLE 1  
RANK ORDER FREQUENCIES OF MMPI TWO-POINT CODES (DARTMOUTH II SAMPLE  
OF 609 COLLEGE FRESHMAN MALES)

Code	N	Percent	Cumulative Percent	Code	N	Percent	Cumulative Percent
35/53	65	10.7	10.7	69/96	7	1.1	88.0
59/95	64	10.5	21.2	68/86	7	1.1	89.1
45/54	41	6.7	27.9	28/82	6	1.0	90.1
25/52	40	6.6	34.5	23/32	6	1.0	91.1
57/75	38	6.2	40.7	29/92	6	1.0	92.1
58/85	35	5.7	46.4	38/83	6	1.0	93.1
56/65	31	5.1	51.5	67/76	6	1.0	94.1
49/94	24	3.9	55.4	24/42	5	.8	94.8
34/43	24	3.9	59.3	14/41	5	.8	95.7
05/50	20	3.3	62.6	02/20	4	.7	96.4
48/84	19	3.1	65.7	06/60	4	.7	97.1
36/63	15	2.5	68.2	17/71	3	.5	97.6
47/74	15	2.5	70.7	18/81	3	.5	98.1
78/87	14	2.3	73.0	16/61	3	.5	98.6
15/51	13	2.1	75.1	09/90	2	.3	98.9
27/72	13	2.1	77.1	19/91	2	.3	99.2
13/31	13	2.1	79.3	12/21	1	.2	99.4
36/63	13	2.1	81.4	26/62	1	.2	99.6
79/97	12	2.0	83.4	04/40	1	.2	99.8
46/64	8	1.3	84.7	01/10	0		99.8
37/73	7	1.1	85.8	03/30	0		99.8
89/98	7	1.1	86.9	07/70	0		99.9
				08/80	0		99.8

tion of 58/85) occurring significantly more frequently among males with education of 13 years or more. The Ss in the present study, as those in the previously reported study (3), however, had completed only 12 years of education at the time of testing. This lends added support to the previous speculation by Gulas (3) that Scale 5 is related more to *IQ* and/or educational aspiration and/or socioeconomic status than to years of education *per se*.

## 2. Code Distribution Comparisons

Chi-square analyses were performed on the code combinations between the two Dartmouth samples. Because of the heightened probability of Type I error given the large number of analyses performed, the alpha level was set at the .01 level.

The statistical comparisons revealed *no* significant differences between the two two-point code distributions. The significance of the result is twofold: (a) it adds validity to the previous clinical interpretations hypothesized by Gulas (3); and (b) it suggests that the reported code distributions are reliable, since two independent samples were utilized and there was a two year time lapse between sampling.

As previously stated (3), caution must be exercised, however, in generalizing these results to college students who differ greatly from the Dartmouth population.

## REFERENCES

1. AARONSON, B. S. Age and sex influence on MMPI peak distribution in an abnormal population. *J. Consult. Psychol.*, 1958, 22, 203-206.
2. GOUGH, H. G. Some personality differences between high ability students who do, and do not, go to college. *Amer. Psychol.*, 1954, 9, 599.
3. GULAS, I. MMPI 2-point codes in normal college populations. Presented at the Ohio Psychological Association, Columbus, October 25, 1972.
4. GUTHRIE, G. M. A study of the personality characteristics associated with disorders encountered by internists. Ph.D. dissertation, University of Minnesota, Minneapolis, 1949.
5. HATHAWAY, S. R., & MEEHL, P. E. An atlas for the clinical use of the MMPI. Minneapolis: Univ. Minnesota Press, 1951.
6. WEBB, J. T. The relation of MMPI two-point codes to age, sex, and education level in a representative nationwide sample of psychiatric outpatients. Presented at the Southeastern Psychological Association, Louisville, Kentucky, April 25, 1970.

*Department of Psychiatry  
Beth Israel Hospital  
330 Brookline Avenue  
Boston, Massachusetts 02215*



## SOME EFFECTS OF FEAR OF FAILURE IN THE ACADEMIC SETTING\*<sup>1</sup>

*Indiana University-Purdue University at Indianapolis*

---

CARLOS GOLDBERG

---

### SUMMARY

This study investigated some of the effects of a fear of failure that is related to self-esteem. The Test Anxiety Questionnaire (TAQ) and the Self-Esteem Contingency Questionnaire (SEC) were administered to the Ss; the SEC measured the extent to which the S's self-esteem is contingent on academic achievement. The Ss were 49 American male college students, and an S was classified as having fear of failure if he scored high on both the SEC and the TAQ.

It was expected that the fear of failure group would have an approach-avoidance conflict about academic pursuits, and measures reflecting this conflict were administered to the Ss.

The results provided partial confirmation for our expectations in that the High SEC group as compared to the Low SEC group had a higher n Ach score, was more compulsive about studying and considered academic degrees as being very important. Furthermore, the fear of failure group (High TAQ-High SEC) tended to have a lower self-esteem and a more negative attitude towards college. The last finding was related to Merton's (14) theory on types of adaptation to society. It was argued that although the fearful of failure individual has a negative attitude towards college, the opposite is not necessarily the case.

---

### A. INTRODUCTION

Atkinson (1, 2) assumes that the motive to achieve success and the motive to avoid failure are both aroused when the individual is confronted with an achievement task. Furthermore, the individual's behavior is a function of the

---

\* Received in the Editorial Office on May 7, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This study was supported by a Grant-in-Aid of Research from Indiana University. Computer service was provided by the IUPUI Computing Facilities. I am grateful to Dr. J. Norton for his assistance with the statistical analyses. I would also like to thank Anne Foster for her able assistance at various stages of the investigation.



difference between the tendency to approach success and the tendency to avoid failure. The individual whose motive to achieve success exceeds the motive to avoid failure is said to have an achievement-oriented personality, whereas if the motive to avoid failure exceeds the motive to achieve, he is said to have a failure-threatened personality (*cf.* 3).

In general, investigators in this area of research conceive of the failure-threatened person as someone who has failed in the past, who stays in the achievement situation because of social constraints, and who obtains a low *n* Ach score and a high score on the Test Anxiety Questionnaire (4, 13).

Birney, Burdick, and Teevan (5) have expanded the concept of the failure-threatened personality by distinguishing different types of fear of failure, one of which is based on the devaluation of the self-estimate that may follow a failure experience. The present study investigated some of the consequences of fear of failure when it is related to the devaluation of the self-estimate.

It is reasonable to assume that some people define themselves primarily in terms of academic achievement; i. e., their self-esteem is contingent on the degree of academic success so that if they do well, their self-esteem is enhanced, but if they do poorly or fail to meet their standards, they feel deflated and insecure.

The individual who fits the above pattern is bound to develop an intense fear of failure, since failure has a deeper connotation for him; viz., it reflects on his self-definition. Thus, unlike the conceptualization of Atkinson *et al.*, it is quite possible that an individual whose motive to achieve success exceeds the motive to avoid failure, who has been relatively successful in the past, may still be fearful of failure.

It can be said, then, that a person whose fear of failure is related to concerns with self-esteem is likely to find himself in a state of approach-avoidance conflict with respect to academic pursuits. On the one hand academic achievement is very important to him because his self-esteem is contingent on success, but on the other hand the academic setting makes him anxious, and he wishes to avoid it altogether.

It is possible, then, that the approach side of this conflict may lead the fearful of failure to engage often in achievement imagery or fantasy, to set unrealistic levels of aspiration, to be compulsive about studying, and to try always to prove to himself and others that he is a highly intelligent and capable individual. In other words, these manifestations serve to bolster the self-esteem of the fearful of failure.

The avoidance side of the conflict should lead the fearful of failure to develop a general negative attitude towards college, courses, examinations,

and other types of evaluation standards. There are two reasons for the last expectation. First, the academic setting produces anxiety in the fearful of failure, and an object associated with anxiety is likely to evoke a negative attitude. Second, a negative attitude towards academic standards serves to protect the fearful of failure should he actually fail; i.e., he can then argue that the established criteria for measuring success are not valid indicators of his intellectual ability (*cf.* 5, p. 213).

Finally, it is expected that the fearful of failure has a low self-esteem, since we assume that only someone with a poor image of himself would feel the need to base his self-estimate primarily in terms of such external criteria as academic success.

## B. METHOD

### 1. Subjects

The subjects were 49 American male students attending Indiana University-Purdue University at Indianapolis. All of the Ss were sophomores, juniors, or seniors, having at least a "B" average in their major. In other words, the Ss had been relatively successful in the past.

### 2. Fear of Failure Measures

Our conceptualization of fear of failure required measures that reflect (a) the extent to which the S's self-esteem is contingent on academic success and (b) the degree of the S's anxiety in the academic setting. Two instruments, described below, were used to measure these two components of fear of failure, and an S was classified as being fearful of failure if he obtained high scores on both measures.

a. *Self-Esteem Contingency Questionnaire* (SEC). The author developed a set of five questions to measure the extent to which one evaluates the self in terms of achievement. The following are the SEC questions:

1. Some students claim that their feelings of self-worth are, in part, dependent on how much they know. As compared with other students, would you say that in your case the connection between knowing and feeling worthy is: Stronger, The Same, or Weaker?

2. Some students report that they feel great when they get a good grade, but when they get a poor grade they feel very low about themselves. To what extent does this apply to you?

3. Some students claim that when they don't know the answer to a question asked of another student, they feel like failures. To what extent does this apply to you?

4. Some students say that when they are reading something and don't

understand it right away, they feel stupid. To what extent does this apply to you?

5. The most important thing for some students is to be very intelligent and knowledgeable. To what extent does this apply to you?

The first question was given the weight of 3, 2, or 1, depending on the S's answer, and each of the last four questions was followed by a six-point rating scale. Thus, the range for the SEC scores is 5-27, where the higher the score the greater the contingency between self-definition and achievement.

Prior to the present study the SEC questionnaire had been administered to 134 American male Ss. A factor analysis yielded only one factor with an eigenvalue greater than one (*cf.* 6). The eigenvalue for this factor was 1.99, and the factor accounted for 40% of the total variance.

*b. Test Anxiety Questionnaire (TAQ).* To measure degree of anxiety in the academic setting, a modified version of the Test Anxiety Questionnaire (TAQ) developed by Mandler and Sarason (12) was administered to the Ss. The TAQ consisted of 12 questions, each of which was followed by a six-point rating scale, where the higher the score the greater the anxiety.

*c. Other measures inappropriate.* It was deemed inappropriate to use the more common measures of fear of failure for the following reasons:

Atkinson's operations (*cf.* 3) could not be used in this study, since low scores on *n* Ach are part of the definition of fear of failure, whereas we expect that our type of fearful of failure will obtain high *n* Ach scores.

The Hostile Press measure of fear of failure developed by Birney, Burdick, and Teevan (5) was also deemed inappropriate. First, Birney *et al.* developed their scoring system on the assumption that Ss who obtain low *n* Ach scores are fearful of failure. In addition, as was argued by Heckhausen (8), the Hostile Press measure may reflect a fear of failure based on concerns with social devaluation rather than self-esteem.

### 3. Other Measures

*a. Need achievement.* The four TAT pictures recommended by Atkinson (1) were presented to the Ss under neutral conditions.

*b. Level of aspiration.* The Ss read a brief excerpt on the neuron taken from an introductory psychology text. They were then asked to predict how well they would do on a quiz testing the material read and were given a quiz which consisted of 10 multiple-choice and true-false questions. In other words, this task can be considered a reading comprehension test.

There were two measures of level of aspiration: (*a*) the S's predicted score yielded the level of aspiration expected score. (*b*) The difference between ob-

tained and predicted score yielded the level of aspiration goal-discrepancy score (*cf.* 10), where the greater the minus score the more the *S* exaggerated his actual ability.

c. *Compulsiveness*. The *S* was asked to answer five questions, each of which was followed by a six-point rating scale: for example; "To what extent do you feel guilty when not studying for your courses?" The higher the score the greater the compulsiveness.

d. *Proving oneself*. There were two questions that measured the extent to which the *S* has a need to demonstrate his intellectual ability, each of which was followed by a six-point rating scale, where the higher the score the greater the need to prove oneself. For example, "How often do you seek out situations where there is an opportunity to demonstrate your intellectual ability?" We also tentatively assumed that the intensity of importance of getting an undergraduate degree and of going to graduate school are indirect measures of the need for proving oneself. Consequently, three questions on the importance of the undergraduate degree and two questions concerning graduate study were administered to the *Ss*. Each of the questions was followed by a six-point rating scale, where the higher the score the greater the importance.

e. *College attitude*. The semantic differential was used to measure attitudes toward college (*cf.* 15). There were 10 concepts, such as *College*, and eight bipolar scales, such as *valuable-worthless*. The higher the score the more negative the attitude towards college.

f. *Self-esteem*. The scale developed by Helmreich *et al.* (9) was used to measure self-esteem. This scale consists of 48 items, each of which is followed by a five-point scale, where the higher the score the greater the self-esteem.

#### 4. Procedure

Two sessions were required for the study. In Session I, the booklet *College Questionnaire* was administered to 159 American male students. The questionnaire contained the semantic differential, the TAQ, and the SEC questions. The student was asked to give background information, including name and telephone number.

The median score on the TAQ was 42; students scoring above the median were classified as High in anxiety, and those scoring below the median were considered Low in anxiety.

The median score on the SEC questionnaire was 17. Students scoring above the median were classified as High in SEC, and those below the median as Low in SEC.



The classification of the students on the TAQ and SEC yielded four groups, with 15 Ss per group. However, some Ss repeatedly failed to report for the second session, and consequently there were 12 Ss in the High TAQ-High SEC group, eight Ss in the High TAQ-Low SEC group, 15 Ss in the Low TAQ-High SEC group, and 14 Ss in the Low TAQ-Low SEC group.

Session II took place two-to-four weeks after the first session, and it consisted of three parts. First, the four TAT pictures were projected on a screen, 20 seconds for each picture, and the Ss were given about eight minutes to write each story. Second, the excerpt on the neuron was projected on the screen for two minutes; the S was then asked to predict how well he would do on the *Recall Quiz*, and the quiz was given. Finally, the *Personal Inventory* was administered, which consisted of questions on compulsiveness, proving oneself, and the self-esteem scale.

A random procedure was followed when the 49 Ss were called to participate in Session II.

### C. RESULTS

#### 1. Need Achievement

Two judges scored the TAT stories. The percentage of agreement was 87%, and the rank-order correlation was .77 ( $p < .01$ ).

Table 1 presents the mean *n Ach* scores for the four groups. By Bartlett's test<sup>2</sup> the assumption of homogeneity of variance was upheld (16, p. 208), and the analysis of variance yielded only a significant main effect due to SEC ( $F = 5.88$ ,  $df = 1/45$ ,  $p < .05$ ). The mean *n Ach* score for the High-SEC group was 3.82, and for the Low-SEC group the mean was .75.

#### 2. Level of Aspiration

None of the results was significant with respect to the level of aspiration expected score. The overall mean for this measure was 72.48 out of a maximum of 100; i.e., the Ss expected to get about a "C" on the quiz.

Table 1 also presents the mean level of aspiration goal-discrepancy scores for the groups. By analysis of variance, the only significant result was the interaction effect ( $F = 5.38$ ,  $df = 1/45$ ,  $p < .05$ ).

#### 3. Compulsiveness

By analysis of variance, the only significant result was the main effect due to SEC ( $F = 4.92$ ,  $df = 1/45$ ,  $p < .05$ ). The mean compulsiveness score for

<sup>2</sup> Bartlett's test was used in the remaining analyses, and in all instances the assumption of homogeneity of variance was upheld.



TABLE 1  
MEAN SCORES OF NEED ACHIEVEMENT, LEVEL OF ASPIRATION, COLLEGE ATTITUDE, AND SELF-ESTEEM OBTAINED BY FOUR GROUPS CLASSIFIED ON THE BASIS OF THEIR SCORES ON THE SELF-ESTEEM CONTINGENCY QUESTIONNAIRE (SEC) AND THE TEST ANXIETY QUESTIONNAIRE (TAQ)

TAQ	SEC	
	High	Low
<i>Need achievement</i>		
High	2.92	.50
Low	4.73	1.00
<i>Level of aspiration/goal discrepancy<sup>a</sup></i>		
High	-5.00	-20.00
Low	-13.33	-2.86
<i>College attitude<sup>b</sup></i>		
High	318.92	276.50
Low	270.60	306.50
<i>Self-esteem</i>		
High	98.58	121.50
Low	113.00	108.21

<sup>a</sup> The higher the minus score the greater the level of aspiration/expected score as compared to actual performance; i.e., the more unrealistic the S is about his ability.

<sup>b</sup> The higher the score the more negative the attitude.

the High-SEC group was 19.25, and that for the Low-SEC group was 16.56; i.e., the former group was more compulsive about studying than the latter one.

#### 4. Proving Oneself

There were no significant results on the direct measure of proving oneself. However, there was a significant main effect due to SEC on the importance of graduate study ( $F = 7.66$ ,  $df = 1/45$ ,  $p < .01$ ). Graduate study was more important for the High-SEC group than it was for the Low-SEC group, the corresponding means being 9.48 and 7.62.

#### 5. College Attitude

Table 1 presents the mean scores on attitudes toward college. By analysis of variance, the only significant result was the interaction effect ( $F = 9.80$ ,  $df = 1/45$ ,  $p < .01$ ). By Tukey's procedure (*cf.* 7), the High TAQ-High SEC group had a more negative attitude towards college than the Low TAQ-High SEC group ( $p < .05$ ). This result partially confirms the expectation that the fear of failure group would develop a negative attitude towards college. However, the other individual comparisons were not significant.

### 6. Self-Esteem

Table 1 presents the means of self-esteem scores. The analysis of variance yielded only an interaction effect ( $F = 6.69$ ,  $df = 1/45$ ,  $p < .05$ ). By Tukey's procedure, the High TAQ-High SEC group had a lower self-esteem than the High TAQ-Low SEC group ( $p < .05$ ). This partially confirms the prediction that the fear of failure group would have a low self-esteem, but the other individual comparisons were not significant.

### D. DISCUSSION

On balance, it can be said that the study provides partial confirmation for our expectations. We assumed that the individual who bases his self-esteem on academic achievement would have a strong approach tendency towards academic pursuits, and the results do show that the High SEC group obtained higher *n* Ach scores, was more compulsive about studying, and considered graduate study more important, as compared to the Low SEC group. However, the fear of failure group (High TAQ-High SEC) did not manifest these approach tendencies. On the other hand, there was a tendency for the High TAQ-High SEC group to have a negative attitude towards college and a lower self-esteem.

The present study did not find any relationship between fear of failure and level of aspiration. However, Mahone (11) did find that fear of failure Ss, as measured by *n* Ach and the Debilitating Anxiety Scale, tended to have an unrealistic level of aspiration. But, as Birney, Burdick and Teevan (5) claim, there are different types of fear of failure, and it is quite possible that one type would set an unrealistic level of aspiration, and the other type would be more realistic. In this regard, it would be interesting to determine how individuals whose fear of failure is mediated by self-esteem and those whose fear of failure is mediated by social devaluation would behave in a risk-taking setting (*cf.* 1, 4).

The finding that the fearful of failure tended to have a negative attitude towards college should not lead to the conclusion that the opposite is also true. After all, there are objective reasons as well as psychodynamic factors for criticizing the university. The same argument applies to the result obtained by Birney *et al.* (5, p. 135) that protesters of the Vietnam war are fearful of failure, as measured by the Hostile Press.

In general, it can be said that there are psychodynamic *and* rational factors that lead one to criticize or reject a social system; but the real issue is what set of factors predominates when the individual evaluates the social system.

This kind of consideration can be used to extend Merton's (14) classification of types of adaptation to society. For example, the rebel who rejects both goals and means of the social system primarily for psychodynamic reasons is likely to engage in types of action that would differ from those actions endorsed by the more rational rebel. Furthermore, the degree of commitment to the alternative social system is bound to be different for these two kinds of rebels.

## REFERENCES

1. ATKINSON, J. W. *Motives in Fantasy, Action, and Society*. New York: Van Nostrand, 1958.
2. ———. *An Introduction to Motivation*. Princeton: Van Nostrand, 1964.
3. ATKINSON, J. W., & FEATHER, N. T. *A Theory of Achievement Motivation*. New York: Wiley, 1966.
4. ATKINSON, J. W., & LITWIN, G. H. Achievement motive and test anxiety conceived as motive to approach success and motives to avoid failure. *J. Abn. & Soc. Psychol.*, 1960, **60**, 52-63.
5. BIRNEY, R. C., BURDICK, H., & TEEVAN, R. C. *Fear of Failure*. New York: Van Nostrand, 1969.
6. HARMAN, H. H. *Modern Factor Analysis*. Chicago: Univ. Chicago Press, 1960.
7. HAYS, W. L. *Statistics for Psychologists*. New York: Holt, Rinehart & Winston, 1963.
8. HECKHAUSEN, H. Achievement motive research: Current problems and some contributions towards a general theory of motivation. In W. J. Arnold (Ed.), *Nebraska Symposium on Motivation*. Lincoln: Univ. Nebraska Press, 1968.
9. HELMREICH, R., ARONSON, E., & LEFAN, J. To err is humanizing—sometimes: Effects of self-esteem, competence, and a pratfall on interpersonal attraction. *J. Personal. & Soc. Psychol.*, 1970, **16**, 259-264.
10. LEWIN, K., DEMBO, T., FESTINGER, L., & SEARS, P. Level of aspiration. In J. McV. Hunt (Ed.), *Personality and the Behavior Disorders*. New York: Ronald, 1944.
11. MAHONE, C. H. Fear of failure and unrealistic vocational aspiration. *J. Abn. & Soc. Psychol.*, 1960, **60**, 253-261.
12. MANDLER, G., & SARASON, S. B. A study of anxiety and learning. *J. Abn. & Soc. Psychol.*, 1952, **47**, 166-173.
13. MCCLELLAND, D., ATKINSON, J. W., CLARK, R. A., & LOWELL, E. L. *The Achievement Motive*. New York: Appleton-Century-Crofts, 1953.
14. MERTON, R. K. *Social Theory and Social Structure* (rev. ed.). New York: Free Press, 1957.
15. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The Measurement of Meaning*. Urbana: Univ. Illinois Press, 1957.
16. WINER, B. J. *Statistical Principles in Experimental Design* (2nd Ed). New York: McGraw-Hill, 1971.

Psychology Department  
 Indiana University-Purdue University  
 1201 East 38 Street  
 Indianapolis, Indiana 46205



## INTELLECTUAL ABILITY AS RELATED TO AGE AND STAGE OF DISEASE IN MUSCULAR DYSTROPHY: A BRIEF NOTE\*

*Fitzsimons Army Medical Center, Denver, Colorado*

F. WILLIAM BLACK

The association of muscular dystrophy and mental retardation has become well documented,<sup>1</sup> with previous research suggesting that upwards of 25% of such patients show intellectual deficits.<sup>2</sup> Duchenne's muscular dystrophy (type MD IIIa) is more frequently associated with retardation than other types.<sup>3</sup> Although current information is inconclusive, there is suggestion of an intellectual decline associated with increasing age and muscle weakness.<sup>1</sup> This study investigated the intelligence test performance of a sample of MD IIIa males.

Subjects were 25 white males with a biopsy confirmed diagnosis of MD IIIa. All were in school, were ambulatory or independent with wheelchair, had not suffered obvious social or educational deprivation as determined by history and observation, and were capable of valid performance on all WISC subtests. A random selection of 10 subjects received EEG studies in addition to the psychometric evaluation. The sample mean age was 9.58 ( $SD = 2.68$ ) and the mean WISC Full Scale  $IQ$  was 82.01 ( $SD = 20.86$ ).

WISC  $IQ$  for the sample ranged from 45 to 128, with 24% of the sample receiving  $IQ$ s within the Borderline to Moderately Retarded ranges. The distribution of intelligence was lower than and significantly different from normal expectancy ( $\chi^2 = 16.63$ ,  $p < .01$ ). The mean Verbal Scale  $IQ$  (84.82,  $SD = 19.11$ ) and Performance Scale  $IQ$  (80.24,  $SD = 22.43$ ) did not differ significantly ( $t = .86$ ). Pearson product-moment correlations provided negative moderate relationships between  $IQ$  and age ( $r_{wy} = -.405$ ,  $p < .05$ ) and stage of disease ( $r_{wy} = -.537$ ,  $p < .01$ ), utilizing the Zellweger and Hansen<sup>3</sup> system for quantifying the degree of muscle involvement.

Four of the 10 EEG subjects had tracings interpreted as abnormal for age,

\* Received in the Editorial Office on May 7, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Dubowitz, V. Intellectual impairment in muscular dystrophy. *Arch. Dis. Child.*, 1965, **40**, 296.

<sup>2</sup> Worden, D. K., & Vignos, P. J. Intellectual function in childhood progressive muscular dystrophy. *Pediatrics*, 1962, **29**, 968.

<sup>3</sup> Zellweger, H., & Hansen, J. W. Psychometric studies in muscular dystrophy, type IIIa (Duchenne). *Devel. Med. & Child Neurol.*, 1967, **9**, 576.



with abnormalities being diffuse in nature. The mean *IQ* of the abnormal EEG subjects was 84.50, which approximates the total sample mean.

Clear moderate negative correlations existed between WISC *IQ* and both age and stage of muscle involvement for this sample. This apparent decline in test performance with increasing age suggests that the degree of retardation in some MD IIIa patients may not be stable over time. Longitudinal studies are necessary for a more definitive explanation of this process. Intellectual deficits were global rather than restricted to verbal or nonverbal performance. An increased incidence of mental retardation, a suggestion of intellectual deterioration with age, and a higher than chance incidence of EEG abnormalities suggest the possibility of an organic etiology for the retardation in MD IIIa. This has previously been suggested in neuropathic studies of such patients.<sup>4</sup>

Although effort was made to exclude subjects with obvious deprivation, the effects of anxiety and depression which must accompany this crippling and fatal disease must be considered in explaining apparent intellectual deterioration.

*Psychology Service*  
*Department of Psychiatry*  
*Fitzsimons Army Medical Center*  
*Denver, Colorado 80240*

---

<sup>4</sup> Rosman, N. P., & Kakulas, B. A. Mental deficiency associated with muscular dystrophy. *Brain*, 1966, **89**, 769.

## THE PERCEPTIONS OF PAROLEES AND PAROLE OFFICERS\* <sup>1</sup>

*Faculty of Business Administration and Commerce,  
The University of Alberta, Canada*

---

EMILE S. SHIHADAH

---

### SUMMARY

This study examined parolees' perceptions of the rehabilitative value of one of the major activities of the parole officers: namely, their investigation of an offender's adjustment to prison life and the assessment made of significant changes in the attitude of the parolee. The theory of delinquent subcultures was used to shed some light on the discrepancy between socially induced aspirations and perceptions of institutional investigation. Dissonance theory was used to account for some of the effects of psychological variables on outcomes of behavior. Degree of favorableness of parolees' perceptions of an institutional investigation was treated as a function of their mode of living: in a halfway house; with parents, relatives, family, wife; or living alone.

It was found that those living with parents, family, relatives, or wife had significantly more favorable perceptions of institutional investigation than those living in a halfway house. A comparison of the perceptions of institutional investigation by parolees and parole officers showed that only those parolees living with parents, family, relatives, or wife had perceptions as favorable as those of the parole officers.

---

### A. INTRODUCTION

Parole systems in Canada and elsewhere are becoming increasingly concerned with the rehabilitation of offenders. Expressions of a general concern are found in public pronouncements, political debates, and formal policy statements of institutions, penitentiaries, and parole departments. Research inquiries have been directed towards the methods used for treatment and their effectiveness and the relative success parole officers have had for one type of offender as compared to another (5, 6, 8, 11). A number of experiments in Britain and the United States have focused on parole supervision and the parolee's expectations from parole activity (4, 7, 9).

---

\* Received in the Editorial Office on May 7, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was partly supported by a research grant from the Department of the Solicitor-General, Ottawa.

The two major parole functions that are used to describe the role of parole officer and the process of interaction with a parolee are investigation and supervision (4, 7). The former refers to the activities of an officer in an institution and the collection of information through his interviews with inmates and staff, together with those inquiries the parole officer undertakes in the community itself where the offender has had some contact and where he will subsequently choose to reside. Supervision takes the form of guidance and control within the constraints of the conditions of parole under which an offender would be released. This study examines and compares the perceptions of parole officers and parolees of the utility of investigating an offender's adjustment while in the institution and his significant changes in attitude (henceforth referred to as "institutional investigation"). This study is part of a much larger investigation of parole officers' and parolees' perceptions of several functions of parole.

In the initial phases of the larger study, the author identified the meaning of rehabilitation for parole officers and parolees in Western Canada. He found that to the parolees it meant "finding a job," "working with people," "having a place to stay," "going straight," "keeping out of trouble." To the parole officers it meant "social adjustment by the parolee to life in the community," "establishing himself as a responsible member of the community," "ability to make wise decisions," "respect for the law." These responses do not necessarily indicate the existence of basic underlying differences in the perceptions of parolees and parole officers of the meaning of rehabilitation. They do suggest, however, that parolees and parole officers saw rehabilitation from different frames of reference. Parolees tended to view rehabilitation in terms of the attainment of immediate, specific, personal goals, while parole officers interpreted it in terms of relatively broad, societally oriented goals. In these circumstances it seemed highly probable that parole officers and parolees would differ sharply in their perceptions of the value of parole activities directed towards the rehabilitation of parolees.

## B. CONCEPTUAL FRAMEWORK

The conditions under which a parolee is released and the ability of a parole officer to change, delete, or modify these conditions create a potential for fear and hostility between the two groups, as well as the need for conformity by the parolee to those conditions and their strict observance. Some studies were concerned with the parolees' anticipation of difficulties in conforming to these conditions (4, 9). Other studies were concerned with the parole officer's role in constructively using his powers to help a parolee adjust to community life

(9). The potential for hostility between the parole officers and parolees stems from the role of parole officers as agents of society for enforcing the conditions of parole, so that irrespective of the declared intentions of the system of parole towards the rehabilitation of parolees, their very activities of rule enforcement make their role, with its punitive consequences, an essential part of the realities of the system of parole. On the other hand the powers of a parole officer to change, delete, or modify the conditions of parole under which a parolee has been released and on which his continued freedom depends provide for an element of reward which is also an essential part of the realities of the system of parole (2, 9). The conceptual framework that guided this study recognized that the conflict inherent in the nature and objectives of parole was likely to result in widely differing perceptions of parole activity by the parole officers and parolees—even of those primarily directed towards the rehabilitation of parolees. The objective of this study was to ascertain with regards to institutional investigation whether in fact such differences in perception did exist, and whether these differences were related to the parolees' choice of residence which formed part of the conditions of release.

The author accepts Cloward and Ohlin's (1) thesis that "core" members are persons who experience a marked discrepancy between socially induced aspirations and the possibilities of achievement. What is not apparent in their assumptions, however, are the conditions under which circumstances on the offender's side interact to determine behavior. On the basis of Cloward and Ohlin's (1) thesis it could be deduced that access to success-goals by illegitimate means may become more apparent depending upon a parolee's mode of living and group associations. Residence in a halfway house, with its relatively greater opportunities of interaction with other parolees, unlike living with parents, family, relatives, or wife, could be more inducive to the favorable perception of illegitimate opportunities. Given the limited access to success-goals by legitimate means that may be induced by one mode of living, the nature of the delinquent response that may result will, therefore, vary according to the availability of various illegitimate means (1). Cloward and Ohlin's (1) propositions, however, do not provide answers to the question of what psychological variables must be taken into account in order to predict outcomes of behavior. Festinger and Aronson's dissonance theory (3) provides some further insights into the interaction of psychological variables and the predictions of the outcomes of various combinations of personality components and situational components. On the basis of some of the major propositions of dissonance theory, parolees would experience dissonant cognitions if they choose one or other alternative form of residence. Furthermore,



once the parolees have decided to adopt a mode of living, they will attempt to convince themselves that the alternative chosen is even more attractive than they had previously thought (3). Those parolees who have selected one or other mode of living will endeavor to reduce their dissonance by (a) enhancing their perceptions of the relative value of the opportunities of one mode of living as compared with another and (b) strengthening the perceived relationship between their present mode of living and the attainment of opportunities (10). The empirical support for these theoretical deductions will now be examined. For that purpose the following hypotheses were formulated:

*Hypothesis 1:* Parolees' perceptions of the value of institutional investigation will be less favorable than those of the parole officers.

*Hypothesis 2:* There will be a positive association between parolees' present mode of living and their perceptions of the value of institutional investigation.

*Hypothesis 3:* The perceptions of parolees who have adopted one or other mode of living will be similar to those of the parole officers with respect to the value of institutional investigation. (Underlying this hypothesis is an assumption that a mode of living is mutually agreed upon between the parole officer and parolee at the time of the parolee's release, and the institutional investigation which has influenced this kind of decision is perceived by the parole officers as contributing positively toward parolees' rehabilitation.)

### C. METHOD

The study was conducted in Western Canada in the Provinces of Alberta and British Columbia. Subjects were 54 parole officers and 132 parolees. This represented 90 percent of the parole officers and 75 percent of the parolees' population.<sup>2</sup> Institutional investigation falls into the following major functions: (a) The collection and collation of factual information through interviews with inmates and staff; (b) the compilation of reports and observations, opinions and assessments, and any other relevant data; (c) the final assessment to the National Parole Board, including recommendations on release, preferences as to mode of living, and a list of general and statutory restrictions on movement within a prescribed radius and such matters as associations with other ex-inmates. The modes of living that were categorized in the larger research study included in a halfway house, with parents, family, relatives or wife, or with no one.

<sup>2</sup> The remainder of the parole officer and parolee populations constituted individuals who were unavailable at the time of the study because of illness or leave of absence. A few persons among the parolees were inaccessible or refused to participate in the study.



Subjects completed a 14 item questionnaire which constituted a Rehabilitative Value Perception Scale (RVPS), specifically designed to measure perception of the rehabilitative value of institutional investigation. Items included in the RVPS were identified by content analysis of interviews with parole officers and parolees. An item was included only if both parole officers and parolees had indicated that it had entered into their evaluations of the utility of institutional investigation in effecting the rehabilitation of parolees. In addition, so as to ensure the unidimensionality of the RVPS, all items had to meet an interitem consistency criterion of a .20 Kendal correlation coefficient with all other items in the scale. Examples of the items included in the RVPS are the following statements: "This kind of information is really the cornerstone on which a parolee's rehabilitation is built." "Whether a parolee is doing well in the institution or not means nothing to his rehabilitation."

Subjects were asked to select one of five alternative responses to each item, indicating the strength of their agreement or disagreement with the item. Choices ranged from "strongly agree," through "neither agree nor disagree" to "strongly disagree." Responses were scored 1 through 5, with a high score indicating a perception of a positive association between a mode of living and rehabilitation of parolees. Scores were normalized according to usual procedures by adjusting for missing observations and dividing by an appropriate factor to obtain scores varying from 0 to 10.

#### D. RESULTS AND DISCUSSION

Hypothesis 1 assumes that parole officers and parolees do not constitute a homogeneous parole population, but represent separate subcultures with differing perceptions of the utility of one or other mode of living. The graphic presentation in Figure 1 of the normalized RVPS scores of the two groups tends to support this assumption. It shows a wider dispersion of parolee scores as compared to those of the parole officers, indicating that parole officers are drawn toward a more uniform evaluation of institutional investigation.

The mean scores of parole officers and parolees were 5.819 and 6.421, respectively. A *t* test indicated that this difference in the mean evaluation by parole officers and parolees of institutional investigation was significant at the .05 level ( $\bar{X}_{Pae} - \bar{X}_{Pa} = -.112$  to  $-1.091$ ). These results support Hypothesis 1.

Hypothesis 2 assumes that an underlying factor in the relatively wider dispersion of parolees' scores, as compared to those of the parole officers, was their differential mode of living. Those residing with parents, family, rela-

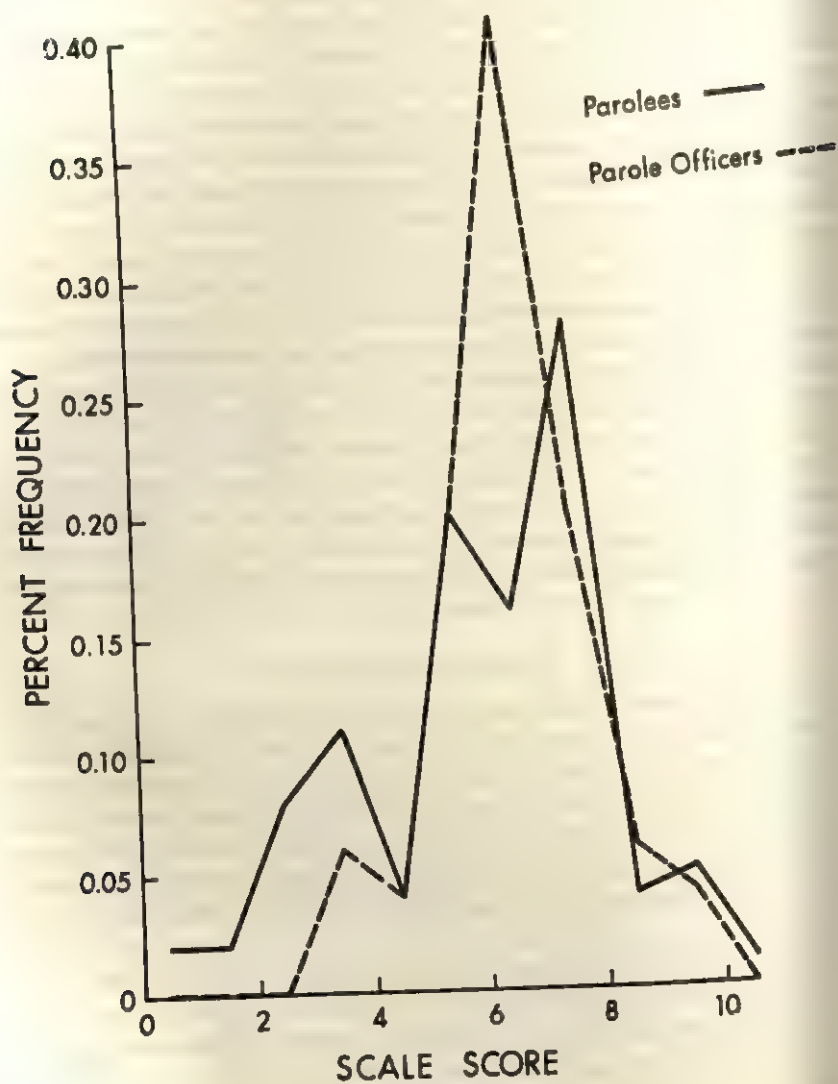


FIGURE 1  
PAROLEES' AND PAROLE OFFICERS' SCORES (NORMALIZED) ON THE REHABILITATIVE  
VALUE PERCEPTION SCALE (RVPS)

TABLE 1  
 t TESTS ON SCORES OF PAROLEES AND PAROLE OFFICERS ON THE REHABILITATIVE  
 VALUE PERCEPTION SCALE (RVPS) BY MODE OF LIVING

Mode of living	Parolees				Parole Officers			
	N	Mean	SD	$\bar{X}_{Pae} - \bar{X}_{Pa}$	N	Mean	SD	$\bar{X}_{Pae} - \bar{X}_{Pa}$
Parents, family, relatives, or wife	18	6.782	1.354		54	6.421	1.236	1.05 to -.32
Halfway house	86	5.575	2.221	-.13 to -2.28*	54	6.421	1.236	-.19 to -1.50*

\* Significant at .05 level.

tives, or wife had significantly more favorable perceptions of the value of institutional investigation than those residing in a halfway house. There were no differences in perception of the value of institutional investigation with those living alone. Thus Hypothesis 2 was supported only with respect to those living with parents, family, relatives, or wife.

From the data presented in Table 1, Hypothesis 3 could be meaningfully tested with respect to those living with parents, family, relatives, or wife, as compared to those residing elsewhere, such as a halfway house. Table 1 indicates that residence with parents, family, relatives, or wife is associated with parolees' perceptions of institutional investigation that are as favorable as those of the parole officers. In the case of those living in a halfway house, parolees maintain perceptions of institutional investigation that are significantly less favorable than those of the parole officers. Hence Hypothesis 3 is supported with respect to living with parents, family, relatives, or wife.

The findings presented in Table 1 are consistent with the assumptions of the formation of subcultures and dissonance theory. Some modes of living, such as in a halfway house, would lead to an increase in interpersonal sensitivity among parolees and to a closer conformity by parolees to the norms of the parolee subculture, rather than to the arousal of dissonance in that group. By its very nature this mode of living emphasizes a high level of interaction among parolees rather than between parolees and parole officers. When the parolee is living with parents, family, relatives, or wife a parole officer is provided with an opportunity for a high level of interaction directly with the parolee or through those with whom he is living which would, in turn, lead to an increase in interpersonal sensitivity between the parole officers and parolees. It is not immediately obvious why there were no significant differences in perceptions between those parolees living alone and those living in a halfway house or with parents, family, relatives, or wife. Presumably the fact of living alone precludes a high level of interaction between those parolees and parole officers or between the parolees themselves. Living alone, in itself, may cause the parolee to consider himself outside of the system and avoid the coercive or sanctioning powers of the parole officer.

#### REFERENCES

1. CLOWARD, R. A., & OHLIN, L. E. *Delinquency and Opportunity: A Theory of Delinquent Gangs*. New York: The Free Press, 1960.
2. CONRAD, J. P. *Crime and Its Correction: An International Survey of Attitudes and Practices*. Los Angeles: Univ. California Press, 1967.
3. FESTINGER, L., & ARONSON, E. Arousal and reduction of dissonance of social contacts. In D. Cartwright & A. Zander (Ed.), *Group Dynamics*. New York: Harper & Row, 1968. Pp. 125-136.

4. GLASER, D. *The Effectiveness of a Prison and Parole System*. New York: Bobbs-Merrill, 1969.
5. GRANT, J. D., & GRANT, M. Q. A group dynamics approach to the treatment of non-conformists in the Navy. *Ann. Amer. Acad. Polit. & Soc. Sci.*, 1959, **322**, 126.
6. HAVEL, J. SIPU—Phase 4: The parole outcome study. Research Report No. 13, Department of Corrections, Sacramento, California, 1965.
7. HOOD, R., & SPARKS, R. *Key Issues in Criminology*. New York: World Univ. Library, 1970.
8. JESNESS, C. F. The Fricot Ranch study. Research Report No. 47, California Youth Authority, Sacramento, 1965.
9. OHLIN, L. E., HERMAN, P., & PAPPENFORTH, D. M. Major dilemmas of the social worker in probation and parole. *Nat. Probab. & Parole Assoc. J.*, 1956, **2**, 211-225.
10. SHIHADAH, E. S., & NEDD, A. N. B. The perceptions of penitentiary inmates and staff. *J. Soc. Psychol.*, in press.
11. SULLIVAN, C. E., GRANT, J. D., & GRANT, M. Q. The development of interpersonal maturity: Applications to delinquency. *Psychiatry*, 1957, **20**, 373.

*Faculty of Business Administration and Commerce*  
*The University of Alberta*  
*Edmonton, Alberta, Canada T6G 2G1*





## CONCEPT-SCALE INTERACTION WITH THE SEMANTIC DIFFERENTIAL TECHNIQUE\*<sup>1</sup>

*Virginia Polytechnic Institute and State University*

DAVID L. KLEMMACK AND JOHN A. BALLWEG

### SUMMARY

Use of the semantic differential technique has largely been restricted to "paper and pencil" settings and student populations. In this study, adult respondents in an interview setting were asked to evaluate two concepts—use of lake (*a*) as a source of electric power and (*b*) as a recreational area—by means of eight seven-point scales. Results indicated considerable concept-scale interaction, suggesting that Osgood's interpretation of the semantic differential may be restricted by both technique of data collection and subpopulation. Respondent hostility might be an intervening variable. The semantic differential technique appeared to yield a preliminary evaluation of the two concepts, suggesting its use as an indicator of orientation.

### A. INTRODUCTION

The semantic differential technique has focused primarily upon the dual questions of the identification of the semantic space necessary to place any given concept and the actual placement of such concepts within a given semantic space (13). This approach, however, ignores the following questions: (*a*) the possibility and consequences of concept by scale interaction (6, 10); (*b*) the effects of intersubject variability in the positioning of concepts in semantic space (3); and (*c*) the methodological issue of whether structural characteristics implied by semantic space theory are due to instrument effects in the semantic differential (1, p. 309). The present study explores the first of these questions, the extent to which a semantic space is confounded by scale-concept interaction. More specifically, we question whether scales used to place concepts in a given semantic space are independent of the concepts that are to be evaluated.

\* Received in the Editorial Office on May 7, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Partial support for this project (Project A-038-Va) was provided by the United States Department of the Interior, Office of Water Resources Research, as authorized under the Water Resources Act of 1964.

Osgood and Suci (13, p. 43) indicate that the bases for the semantic differential are as follows: (a) The process of judging or describing a given concept can be conceived of as a continuum, the ends of which are defined by a pair of polar adjectives. (b) Although meaning can vary in many ways, at least some of the types of meaning are essentially equivalent and thus can be described in terms of a single dimension. (c) A limited number of these dimensions can serve to define a semantic space within which the meaning of any concept can be located. Thus, bipolar adjective pairs may be grouped in terms of a relatively small number of dimensions, a semantic space. The meaning of any concept, in turn, is given by its placement in relationship to the dimensions isolated in the first phase.

Two related approaches have been followed in determining the dimensionality of a semantic space. Both methods begin by having respondents evaluate a number of concepts in terms of a number of scales (bipolar adjective pairs). At this point, however, the two approaches differ somewhat. In the first, the mean or median score for each scale is obtained, and the resulting summary scores are correlated with one another across concepts. The correlation matrix (order  $m$  where  $m$  is the number of different scales employed) is then factored with use of either a centroid or principal factor solution, and the number of dimensions thus isolated is deemed to determine the structure of the semantic space. In the second approach, individual scores are not summarized. Instead, correlations are obtained for each scale associated with each concept across individuals, yielding a substantially larger matrix (order  $mn$  where  $m$  is, as before, the number of unique scales, and  $n$  is the number of concepts evaluated). Again the resulting matrix is factored, but in this case the dimensions isolated may be examined for scale by concept interaction. Such an analysis is, of course, impossible with the first approach.

Although many investigations (4, 5, 11, 12, 14), primarily using mean scores as the basic data, report findings that are generally consistent with the theory underlying the semantic differential, others (3, 10) note that the use of the semantic differential poses both theoretical and methodological problems.<sup>2</sup> Each notes and presents some evidence to indicate that it is, at least occasionally, reasonable to assume that the relationship between any two scales may, in part, be a function of the concepts that are being evaluated. Thus, to use Nunnally's example (10, p. 539), we would expect a positive correlation between a tough-tender and a valuable-worthless scale when evalu-

<sup>2</sup> Although Osgood and Suci (12, 13) consider the problems of both scale-concept and subject-concept placement interaction using a variance source solution as proposed in this study, many followers tend to use the dual source solution. Such an approach is adequate only after the absence of scale-concept interaction has been verified.

ating men, but a negative correlation when evaluating steaks. Although Nunnally does indicate that such interaction is less likely when the concepts to be evaluated are similar, the problem still remains. Theoretically such findings are inconsistent with the framework presented by Osgood in that they imply that the concepts belong, at best, in two different semantic spaces. Pragmatically, concept-scale interaction may be interpreted as indicating that the individuals involved use different bases for the evaluation of the concepts. In either case, the number of dimensions necessary to define the semantic space is increased by interaction, and no simple interpretation of the space is possible.

The present study explores the question of the semantic space necessary to define two related concepts: the use of a lake as a source of electric power and the use of the same lake as a recreational area. Specifically, the question posed is this: "May the two concepts be placed in the same semantic space, or are the responses to the semantic scales concept specific?"

## B. SAMPLE AND METHOD

While the use of the semantic differential has generally been restricted to a questionnaire approach, the data for the present study were collected by means of interviews. The sample consisted of 307 adult respondents, one per household, where households were selected by a cluster technique (8, p. 48) with use of 1970 Census tracts for Roanoke, Virginia.<sup>3</sup> The interview schedule contained, imbedded among other questions, two identical sets of eight seven-point semantic differential scales: good-bad, beautiful-ugly, valuable-worthless, pleasant-unpleasant, happy-sad, fair-unfair, useful-useless, and important-unimportant. In an effort to avoid response bias in the form of a "halo effect," the order of every other adjective pair was reversed. Prior to the administration of the semantic differential, it was ascertained that all subjects were aware of a man-made lake in the area, known as Smith Mountain Lake. The lake serves both as a recreational area for inhabitants of the city and as a source of hydroelectric power. The interviewee was presented with a card listing the eight seven-point adjective scales and asked to respond to each of the scales in relation to the statement: "Consider Smith Mountain Lake as a recreational area." Respondents were then requested to respond,

<sup>3</sup> Attempts were made to reach adults in 370 different households. In 53 cases either all eligible respondents at home at the time of contact refused to be interviewed, or no eligible respondents were available at the time of contact. The 10 remaining cases excluded consisted of interview schedules that, for one reason or another, were incomplete, leaving a final *N* of 307.

using the same eight scales, to the statement: "Consider Smith Mountain Lake as a source of electric power."

Since the initial question posed in the study focuses on the possibility of concept by scale interaction, the distinction between the two statements was retained. Responses to each of the 16 scales (eight for each statement) were correlated, and the resulting matrix factored by use of a principal factor solution with squared multiple correlations as initial communality estimates. Initial estimates as to the completeness of factorization were made according to Kaiser's criterion (7, p. 198), and, as a check on the adequacy of the initial solution, the corresponding residual matrix examined. After ascertaining the "appropriate" solution in terms of dimensionality, the factor matrix was rotated to "simple structure form" according to the varimax criterion (7, pp. 305-306). If concept by scale interaction was not present, it would be expected that identical scales applied to two different concepts would have high factor loadings on the same factor. If, on the other hand, there was considerable concept by scale interaction, we would expect that scales associated with one concept would load on one factor, while those associated with a second concept would load on another factor.

### C. RESULTS

Examination of the nine positive eigenvalues obtained suggests that the appropriate factor solution consists of two dimensions. The eigenvalues associated with the first and second factors were substantially above one (5.1 and 2.1, respectively), while that for the third factor was somewhat below one (.67), implying that at least two but probably not three factors are necessary to reproduce the initial correlation matrix adequately. At the same time, the use of Kaiser's criterion for the completeness of factorization is inadequate unless the estimates of the item communalities approach one (7, p. 198), a condition that is clearly not met in these data (see Table 1). While there is little doubt that at least two factors are necessary, a more complex solution may be more appropriate.

As a further aid in the selection of the most appropriate factor matrix, a residual matrix with two factors extracted was obtained. Examination of the off diagonal elements suggested no pattern in the residual correlations, adding further support to the hypothesis that a two factor solution was sufficient. Next, the communality estimates obtained from the two factor solution were compared to the initial estimates. The two sets were similar (absolute mean difference of .04 with a maximum deviation of .08), implying that the two factor solution did account for most of the common variance. Finally, a



TABLE 1  
FACTOR LOADINGS ON TWO CONCEPTS OF EIGHT SEMANTIC DIFFERENTIAL SCALES

Concept-scale	Factor 1	Factor 2	$h^2$
Recreation			
valuable-worthless	+ .002	.725	.526
fair-unfair <sup>a</sup>	-.114	.510	.273
useful-useless	-.087	.713	.516
pleasant-unpleasant <sup>a</sup>	-.158	.535	.311
happy-sad	-.266	.637	.477
important-unimportant <sup>a</sup>	-.097	.668	.456
good-bad	-.234	.725	.580
beautiful-ugly <sup>a</sup>	-.344	.370	.255
Electric Power			
valuable-worthless	-.686	.137	.489
fair-unfair <sup>a</sup>	-.600	.087	.368
useful-useless	-.626	.229	.444
pleasant-unpleasant <sup>a</sup>	-.731	.133	.535
happy-sad	-.745	.092	.563
important-unimportant <sup>a</sup>	-.602	.150	.385
good-bad	-.784	.091	.623
beautiful-ugly <sup>a</sup>	-.584	.191	.378
% of trace	44.26	36.81	81.07
% of total variance	24.55	20.47	45.02

<sup>a</sup> Pairs reversed to minimize response bias.

three factor solution was obtained and rotated, but efforts at interpretation proved fruitless, thus adding a further modicum of support to the hypothesis that a two factor solution is sufficient.

Examination of the two factor rotated matrix (see Table 1) indicates that there was considerable question by scale interaction. All eight of the adjective pair scales associated with the question concerning the use of the lake as a source of electric power had substantial loadings on the first factor and minimal loadings on the second. Furthermore, seven of the eight scales associated with the question concerning the use of the lake as a recreational area loaded primarily on the second factor with only minimal loadings on the first. Thus, judgment of the two uses of the lake, recreational and power, appeared to be independent, with the scales forming two separate evaluative dimensions, one for each type of use. Such an interpretation is further supported by an examination of the one scale, beautiful-ugly in response to the recreation question, that loaded on both dimensions or factors. The beauty of the lake as a recreational area could indeed be harmed by the parallel use of the lake as a source of electric power. We might then expect that the judgment of the lake as a recreational area in terms of the beautiful-ugly scale would be confounded with the judgment of the lake as a source of electric power, an observation consistent with the findings.

To summarize briefly, respondents appear to have used the same set of semantic differential scales to make separate judgments of the use of lake as (a) a source of electric power and as (b) a recreational area. The two types of judgments, however, appear to have been made in opposite directions. The use of the lake as a recreational area tended to be judged along as positive continuum, while the use of the lake as a source of power tended to be judged along a negative continuum. In any case, the data show considerable concept-scale interaction in that the same set of scales loaded on different dimensions dependent upon the concept to which they were applied.

#### D. DISCUSSION

The results clearly indicate that, although the concepts to be evaluated appear to be related, there was considerable concept by scale interaction. We would normally expect that the possibility of this type of interaction is minimal, but, as these data suggest, minimal does not imply impossible. Rather than obtaining a differential location of two concepts in a single semantic space as hypothesized by Osgood, we found that the concepts themselves are dimensionally different. The judgments about the lake as a recreational area were, if we are to believe the data, made according to different criteria than were the judgments about the lake as a source of electric power. Thus we might tentatively conclude that it is indeed possible to find scale by concept interaction even with related questions or concepts, suggesting, as an initial implication of the results of this study, that it is always wise to consider explicitly whether a given set of semantic differential data is confounded by question-scale interaction.

Although there are many possible explanations for the failure of the data to be consistent with Osgood's framework, one of the most apparent lies in the nature of the data collection process used. In most cases, the semantic differential has been administered in a "paper and pencil" setting, but we used an interview mode for gathering our data. We could hypothesize that at least part of the general consistency in the findings of others when using the semantic differential is a function of the testing format used. It could well be that responses to paper and pencil tests are, even with identical questions, different from those obtained with use of an interview procedure. Certainly, the language skills necessary to complete the task are different in the two settings. Also, social skills are, without question, much more critical in the second setting. Face to face interaction, even when the task is relatively abstract and neutral, may indeed lead to a differential response pattern on the part of the interviewee. Perhaps, since the task is relatively abstract, it is

more difficult to complete in an interview setting, thus generating a certain amount of hostility on the part of the respondents that would not be present in a paper and pencil setting.<sup>4</sup>

A second possible explanation for the failure of these data to be consistent with Osgood's framework lies in the nature of the population studied. Typically, the semantic differential is not only administered in a paper and pencil setting, but also the subjects tend to be students. We, on the other hand, attempted to apply the technique to a representative sample of adults in a medium-sized city. Thus, we could also hypothesize that at least a part of the consistency of results obtained by Osgood and others is due to the uniform nature of the populations that they have studied. It could be that Osgood's model is appropriate for the subpopulation of college students but not for adults in general. Obviously, students as a category within the general population differ in a variety of characteristics from the population at large, not the least of which is in terms of verbal skills. There is a possibility that such differences lead to a differential response set in the two types of populations that, in turn, accounts for the inapplicability of Osgood's framework for the data gathered in this study. In any case, if either population differences or differential testing situations account for the departure of our results from those hypothesized by Osgood, we would be forced to reconsider the validity of Osgood's model.

In spite of the lack of support for Osgood's model, the results do carry some positive implications. First, if a researcher is simply interested in a positive to negative evaluation of some object, institution, or situation, it appears that the semantic differential scales can provide a fairly efficient means to that end. Certainly, it is faster to hand a respondent a card with eight adjective pairs, present the object to be evaluated, and obtain eight responses than it is to ask the respondent eight separate questions. And if our interpretation of the results is correct, the data resulting from such a procedure represent a crude evaluation of the object or concept in positive to negative terms. The question remaining, of course, is whether the researcher desires such information. Such an approach to the semantic differential has, however, been of value to some investigators [Brinton (2) and, to a lesser extent, Mindak (9)]. If the length of the interview is a critical factor, and if the researcher feels that the potential hostility created by using such an abstract task as the semantic differential is not insurmountable, then the use of the

---

<sup>4</sup> Interviewers reported informally that respondents reacted negatively to the semantic differential component of the questionnaire, thus lending a modicum of support to this hypothesis.

semantic differential scales as a preliminary indicator of attitudes might represent a reasonable alternative to the development of a specific attitudinal scale.

## REFERENCES

1. ANDERSON, A. B. Structure of semantic space. In E. Borgatta & G. Bohrnstedt, *Sociological Methodology*. San Francisco: Jossey-Bass, 1970. Pp. 308-325.
2. BRINTON, E. Deriving an attitude score from SD data. *Public Opin. Quart.*, 1961, **25**, 289-295.
3. CARROLL, J. B. The Measurement of Meaning, by C. E. Osgood, G. T. Suci, and P. H. Tannenbaum: A book review. *Language*, 1959, **35**, 58-77.
4. DEESE, J. On the structure of associative meaning. *Psychol. Rev.*, 1962, **69**, 161-175.
5. FLAVELL, J. H. Meaning and meaning similarity: I. A theoretical reassessment. *J. Gen. Psychol.*, 1961, **64**, 307-319.
6. GULLIKSEN, H. How to make meaning more meaningful. *Contemp. Psychol.*, 1958, **3**, 115-118.
7. HARMON, H. H. *Modern Factor Analysis*. Chicago: Univ. Chicago Press, 1967.
8. KISH, L. *Survey Sampling*. New York: Wiley, 1967.
9. MINDAK, W. A. Fitting the sample differential to the marketing problem. *J. Market.*, 1961, **25**, 28-33.
10. NUNNALLY, J. *Psychometric Theory*. New York: McGraw-Hill, 1967.
11. OSGOOD, C. The cross-cultural generality of visual-verbal synesthetic tendencies. *Behav. Sci.*, 1959, **5**, 146-169.
12. OSGOOD, C., & SUCI, G. Factor analysis of meaning. *J. Exper. Psychol.*, 1955, **50**, 325-338.
13. OSGOOD, C., & SUCI, G. Factor analysis of meaning. In J. G. Snider & C. E. Osgood (Eds.), *Semantic Differential Technique*. Chicago: Aldine, 1969. Pp. 42-55.
14. SUCI, G. J. Comparison of semantic structures in American Southwest culture groups. *J. Abn & Soc. Psychol.*, 1960, **61**, 25-30.

*Department of Sociology*

*Virginia Polytechnic Institute and State University*

*Blacksburg, Virginia 24061*



## EFFECTS OF *IQ* AND MENTAL AGE ON VERBAL IMITATIVE PERFORMANCE OF CHILDREN\*<sup>1</sup>

*The University of Georgia*

REX FOREHAND, BRUCE ROBBINS, AND CHARLES PAT BRADY

### SUMMARY

The present study examined the verbal imitative performance of normal and retarded children at four mental ages (*MA*s 5, 6, 7, and 8 years). Three dependent measures of imitation were taken: total, mimical, and conceptual. The results indicated that both *IQ* and mental age were significant factors in imitative performance. Retardates primarily demonstrated mimical imitation at the *MA* 5 level and conceptual imitation at the *MA* 7 level, while normals emitted primarily mimical responses at *MA* 7. Retardates and normals differed in total imitation only at *MA* 5.

### A. INTRODUCTION

Bandura (1), among others, has pointed out the importance of imitation in the learning of children. Most of the research in imitative performance has involved motor tasks. Rickard and his associates (4, 5, 6) have recently used a methodological approach to the study of verbal imitative behavior in which a subject listens to a tape recorded model say words and, in response, the subject emits a word. The data are examined in terms of the number of imitations of a critical word class; i.e., animal nouns. Although most of the Rickard studies involve college students, one recent investigation (7) used kindergarten children. This study found only an increase in critical response words by the subjects as the frequency of such words emitted by the model increased.

In the studies by Rickard and his associates an imitative response was any animal word emitted by the subject. Holt (3) has suggested that the word modeling responses can be divided into two categories: mimical and concep-

\* Received in the Editorial Office on May 7, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This study was supported by Grant MH22403-01 from the National Institute of Mental Health. Appreciation is expressed to the following people and agencies for making the study possible: Dr. Charles Smock and his Follow Through staff, Pickens County (Georgia) School System, and the Georgia Retardation Center. The research was conducted while the second and third authors were predoctoral fellows supported in part by United States Health Service Grant MH08924. Reprint requests should be directed to Rex Forehand at the address shown at the end of this article.



tual. The first type of response refers to the subject emitting the same word as the model. The second type of response, conceptual imitation, refers to the subject emitting a different word from the model but one in the same response class.

Within the area of mental retardation, Turnure and Zigler (8) have found that retarded children imitate a model significantly more than normals on a motor task. Zigler (9) accounts for such a finding in terms of failure and outer-directedness. Retardates experience a higher degree of failure than normals and, consequently, rely less on themselves and more on external cues from others to solve problems. As a result, retardates are more likely to imitate a model than normals. Although Zigler and his associates have examined the effects of numerous factors (9) on imitative behavior of normals and retardates, their research has not typically examined either verbal imitation or the effects of developmental changes on imitation.

The purpose of the present study was to examine changes in verbal imitative performance of retarded and normal children as a function of mental age. It was predicted that both normals and retardates, although perhaps at different rates, would show a decrease in mimical imitation responses and an increase in conceptual imitation with increases in *MA*. This prediction is based on the assumption that conceptual imitation involves a higher level of cognitive functioning than mimical imitation. The latter involves merely the repetition of a modeled word, whereas the former involves formulating the modeled word as a member of a response class and, subsequently, emitting another word within that class. The *MA*s examined were five, six, seven, and eight, since a recent study (2) found that mimic responding was almost totally absent in retardates at the nine year *MA* level. It was also predicted, according to Zigler's theory (9), that retardates would demonstrate more total imitation than normals.

## B. METHOD

### 1. Subjects

The subjects were 32 normals from a public school system and a kindergarten and 32 noninstitutionalized retardates. Eight retardates, four males and four females, were selected from each of four developmental levels: *MA*s 5, 6, 7, and 8. Within each developmental level the retarded subjects were randomly selected from within an *IQ* range of 50 to 80, with the restriction that the subjects had not been diagnosed as having brain damage and were free of

gross motor and sensory deficits. Mental ages and IQs were determined by the Stanford-Binet and WISC.

The 32 normals, four males and four females at *MA* 5, 6, 7, and 8, respectively, were selected to approximate the *MA*s of the retardates. Mental ages and IQs were determined from the California Test of Mental Maturity. The normals were from a rural community and could generally be classified as upper-lower and middle class children. Table 1 presents the characteristics of the subjects.

## 2. Apparatus

The apparatus was a cassette tape recorder and a cassette tape. The stimulus tape was prepared with the voice of a 28-year-old female who was not known by the subjects. The taped list consisted of 50 words. The first 10 words contained no animal words, the critical response class. The next 10 words contained two animals words, the next 10 contained four, the next 10 contained six, and the last 10 contained eight.

## 3. Procedure

After the subject was brought to the experimental room and seated in a chair, the experimenter gave the following instructions: "I want you to listen to this tape player. When you hear a voice say a word, you are to say a word. Say just one word. Say the first word you think of." Immediately following the instructions, the experimenter activated the tape player and the 50 stimulus words were presented at 5 second intervals. The experimenter recorded the subject's responses. At the conclusion of the tape, the experimenter assured the subject that he had "done a good job," thanked him, and answered any questions.

TABLE 1  
MEAN *CA*, *MA*, AND *IQ* FOR DIFFERENT SUBJECT GROUPS

Groups	<i>CA</i> (in months)		<i>MA</i> (months)		<i>IQ</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Retarded						
<i>MA</i> 5	105.4	20.5	64.4	3.0	61.1	6.9
<i>MA</i> 6	124.1	18.5	77.1	3.1	63.5	5.2
<i>MA</i> 7	141.5	20.6	89.9	2.7	64.5	7.3
<i>MA</i> 8	167.2	28.1	100.4	4.9	65.5	9.8
Normal						
<i>MA</i> 5	67.5	4.4	69.5	6.3	102.4	11.7
<i>MA</i> 6	77.9	3.6	79.2	3.1	102.0	7.3
<i>MA</i> 7	89.6	6.1	90.8	3.7	100.0	6.6
<i>MA</i> 8	100.8	4.2	104.2	4.3	101.6	5.0

## C. RESULTS

The data consisted of total imitations, conceptual imitations, and mimical imitations of words in the critical response class: i.e., animal words. Total imitations were defined as the total number of animal responses emitted by the subject. Mimical imitations were defined as emissions by the subject of the same critical class words as the taped model emitted. Conceptual imitations were defined as the emission of a nonmimical animal noun by the subject. Each of the three types of imitation was examined by a  $5 \times 2 \times 2 \times 4$  analysis of variance with one within factor (Trial Blocks) and three between factors (*IQ* level, Sex, and *MA* level). Table 2 presents the mean number of total, conceptual, and mimical imitations at each *MA* level for normals and retardates.

The analysis of total imitations resulted in a significant Trial Block main effect ( $F = 99.92$ ,  $df = 4/192$ ,  $p < .001$ ),  $IQ \times MA$  interaction ( $F = 3.80$ ,  $df = 3/48$ ,  $p < .025$ ), and Trial Block  $\times IQ \times MA$  interaction ( $F = 2.90$ ,  $df = 12/192$ ,  $p < .005$ ). The Trial Block effect resulted from an increasing number of imitations across trial blocks (TB 1 < TB 5,  $t = 92.1$ ,  $df = 63$ ,  $p < .001$ ).<sup>2</sup> The  $IQ \times MA$  interaction (see Table 2) resulted from retardates imitating significantly more than normals at *MA* 5 ( $t = 3.57$ ,  $df = 14$ ,  $p < .01$ ). The Trial Block  $\times IQ \times MA$  interaction indicated that the significantly greater imitation by retardates than normals at *MA* 5 occurred in Trial Blocks 3, 4, and 5 ( $t > 2.82$  in all cases,  $df = 14$ ,  $p < .02$ ).

The analysis of conceptual imitations revealed a Trial Block main effect ( $F = 10.20$ ,  $df = 4/192$ ,  $p < .001$ ),  $IQ \times MA$  interaction ( $F = 2.86$ ,  $df = 3/48$ ,  $p < .05$ ), and Trial Block  $\times IQ \times Sex$  interaction  $F = 2.52$ ,  $df = 4/192$ ,  $p < .05$ ). The Trial Block effect resulted from an increasing number of conceptual imitations across trial blocks (TB 1 < TB 5,  $t = 33.8$ ,  $df = 63$ ,  $p < .001$ ). The  $IQ \times MA$  interaction, as seen in Table 2, resulted from retardates emitting significantly more conceptual imitations than normals at the *MA* 7 level ( $t = 2.86$ ,  $df = 14$ ,  $p < .02$ ) and from retardates emitting more conceptual imitations at the *MA* 7 level than at the *MA* 5 level ( $t = 2.28$ ,  $df = 7$ ,  $p < .06$ ). The Trial Block  $\times IQ \times Sex$  interaction resulted from a lower emission of conceptual animal words by normal females in Trial Block 1 than by normal males and retarded males and females ( $t > 2.042$  in all cases,  $df = 30$ ,  $p < .05$ ).

The analysis of mimical imitations revealed a significant Trial Block main effect ( $F = 69.08$ ,  $df = 4/192$ ,  $p < .001$ ),  $IQ \times MA$  interaction ( $F = 5.61$ ,

<sup>2</sup> All  $t$  probability values are based on two tail tests.

$df = 3/48$ ,  $p < .005$ ), and Trial Block  $\times IQ \times MA$  interaction ( $F = 4.98$ ,  $df = 12/192$ ,  $p < .001$ ). The Trial Block effect resulted from an increasing number of mimical imitations across trial blocks (TB 1 < TB 5,  $t = 63.9$ ,  $df = 63$ ,  $p < .001$ ). An examination of the  $IQ \times MA$  interaction (see Table 2) indicated that normals gave more mimical responses at the MA 7 level than at the MA 5 level ( $t = 2.33$ ,  $df = 7$ ,  $p < .05$ ). Furthermore, significantly more mimical responses were emitted by retardates than normals at MA 5 ( $t = 4.45$ ,  $df = 14$ ,  $p < .001$ ) and significantly more mimical responses by normals than retardates at MA 7 ( $t = 3.21$ ,  $df = 14$ ,  $p < .01$ ). The Trial Block  $\times IQ \times MA$  interaction indicated that the difference in retardates and normals at MA 5 and MA 7 resulted from differences at Trial Blocks 2 through 5 ( $t > 2.81$  in all cases,  $df = 14$ ,  $p < .02$ ).

#### D. DISCUSSION

The results of the present study indicate that differences in normals and retardates in verbal imitative behavior are dependent on both mental age and type of verbal imitation examined. At the MA 5 level retardates demonstrated significantly more mimical imitation and total imitation than normals. At the MA 7 level retardates displayed more conceptual imitation, while normals manifested more mimical imitation. Normals and retardates did not differ in total, conceptual, or mimical imitations at the MA 6 and MA 8 levels.

It was predicted that both normals and retardates would demonstrate a decrease in mimical imitation and an increase in conceptual imitation with increases in MA. From Table 2 it is evident that retardates generally supported the prediction. In contrast, normals increased in mimical imitations from the MA 5 level to the MA 7 level. From an examination of the data of normals at the MA 5 level, it appeared that the subjects possibly failed to comprehend the instructions. In general, many of the responses of normals at MA 5 appeared unrelated to the stimulus words. For example, some subjects

TABLE 2  
MEAN NUMBER OF THREE TYPES OF IMITATION AT EACH MA FOR  
NORMALS AND RETARDATES

Imitation	Mental age							
	Normals				Retardates			
	5	6	7	8	5	6	7	8
Total	7.8	14.1	17.7	13.4	17.2	16.9	14.0	12.4
Conceptual	4.8	4.1	.8	6.0	1.1	6.8	9.1	4.7
Mimic	3.0	10.0	16.9	7.4	16.8	10.1	4.9	7.7



counted in response to the modeled words, while others often failed to respond verbally to a stimulus word.

The finding that retardates gave more total imitations than normals only at the *MA* 5 level does not support Zigler's hypothesis (8, 9) that retardates are more outer-directed and, consequently, imitate more than normals. The present results indicate that with a verbal task the particular *MA* level is a significant factor in comparing imitation of normals and retardates.

The finding of a significant trial block effect supports the previous findings reported by Rickard and his associates (4, 6, 7). Such a finding essentially indicates that with an increasing number of critical words to imitate, subjects demonstrate an increase in imitative responses. Since the studies by Rickard and his associates have examined only total imitations, the initial low rate of emission of conceptual responses by females in the present study cannot be compared to previous research. However, future research should be attuned to possible sex differences.

#### REFERENCES

1. BANDURA, A. *Principles of Behavior Modification*. New York: Holt, Rinehart & Winston, 1969.
2. GARDNER, H. L., & FOREHAND, R. The effects of institutionalization on word responses of retardates. Unpublished manuscript, The University of Georgia, Athens, 1972.
3. HOLT, M. M. Personal communication, 1972.
4. RICKARD, H. C., & JOUBERT, C. Subject-model sexual status and observer performance. *Psychon. Sci.*, 1968, **10**, 407-408.
5. RICKARD, H. C., & WILLARD, D. L. Note on recorded vs. "actual" model behavior and observer performance. *Psychol. Rep.*, 1968, **22**, 839-840.
6. RICKARD, H. C., HOLT, M., & LATTAL, K. A. Reliability of word-modeling performance. *Psychol. Rep.*, 1970, **26**, 550.
7. RICKARD, H. C., ELLIS, N. R., BARNHART, S., & HOLT, M. Subject-model status and verbal imitative performance in kindergarten children. *Devel. Psychol.*, 1970, **3**, 405.
8. TURNURE, J., & ZIGLER, E. Outer-directedness in the problem solving of normal and retarded children. *J. Abn. & Soc. Psychol.*, 1964, **69**, 427-436.
9. ZIGLER, E. Research on personality structure in the retardate. In N. Ellis (Ed.), *International Review of Research in Retardation* (Vol. 1). New York: Academic, 1966.

*Department of Psychology*  
*The University of Georgia*  
*Athens, Georgia 30601*



(Founded in 1935 by Carl Murchison)

# The Journal of PSYCHOLOGY

The General Field of Psychology

POWELL MURCHISON

*Managing Editor*

HELEN M. DAVIS

*Associate Editor*

## *Editorial Board*

JOHN A. ANTOINETTI  
Life Insurance Agency  
Management Association

S. HOWARD BARTLEY  
Michigan State University

ARTHUR G. BILLS  
University of Cincinnati

E. A. BOTT  
University of Toronto

LEONARD CARMICHAEL  
National Geographic Society

IRVIN L. CHILD  
Yale University

HERBERT S. CONRAD  
U. S. Office of Education

JOHN F. DASHIELL  
Florida State University

LEONARD W. DOOB  
Yale University

If this space should be unstamped, this is the regular library edition.  
But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not  
be resold for a period of five years from date of publication.

FRANK A. GELDARD  
Princeton University

ERNEST R. HILGARD  
Stanford University

JOHN E. HORROCKS  
Ohio State University

DAVID HOTHERSALL  
Ohio State University

HEINRICH KLÜVER  
University of Chicago

NORMAN L. MUNN  
University of Adelaide,  
Australia

WARREN H. TEICHNER  
New Mexico State  
University

LEE EDWARD TRAVIS  
Fuller Theological  
Seminary

JOSEPH ZUBIN  
Psychiatric Institute,  
Columbia University

## VOLUME 84

1973

Copyright © 1973 by the Journal Press  
Provincetown, Massachusetts, U. S. A.

All rights reserved

Published at 2 Commercial Street, Provincetown, Mass.

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

# AUTHOR INDEX

Abramson, H. A. ....	111	Hartje Jack C. ....	61
Aliotti, Nicholas C. ....	137	Horrocks, John E. ....	241
Arnold, Thomas C. ....	129	Julian, J. W. ....	13
Balla, David ....	97	Kennedy, Elizabeth ....	75
Baliweg, John A. ....	345	Klein, Robert E. ....	219
Baumeister, Alfred A. ....	267	Klemmack, David L. ....	345
Black, F. William ....	333	Kokonis, Nicholas D. ....	211
Blanton, William E. ....	137	Lerner, Richard M. ....	229
Borman, Walter C. ....	55	Levenson, Hanna ....	237
Boshier, Roger ....	45	Love, Craig ....	253
Brady, Charles Pat ....	353	Lurey, Edward ....	195
Burgess, Gary D. ....	317	Mahoney, E. R. ....	223
Busse, Thomas V. ....	253	McCormick, Mark ....	97
Campbell, Cecil O. ....	267	McGhee, Paul E. ....	189
Carone, P. A. ....	111	Michielutte, Robert ....	33
Chase, Lawrence J. ....	225	Millett, Ricardo ....	219
Chen, Kathleen ....	89	Mills, Norbert H. ....	225
Cochrane, Carl M. ....	33	Mussman, Milton C. ....	241
Cohen, Arie ....	133	Osborne, R. T. ....	159
Dalton, Starrette ....	257	Papalia, Diane E. ....	75
Dwyer, Francis M. ....	129	Peretti, Peter O. ....	81
Farley, Frank H. ....	133	Pirojnikoff, Leo A. ....	105
Fodor, Eugene ....	181	Pollio, Howard R. ....	173
Forehand, Rex ....	353	Robbins, Bruce ....	353
Forthman, John H. ....	23	Rosenbaum, Leonard L. ....	165
Freeman, Howard E. ....	219	Rosenbaum, William B. ....	165
Gavurin, Edward I. ....	279	Ryckman, Richard M. ....	317
Gengerelli, J. A. ....	291	Scharff, William H. ....	283
Gettner, H. H. ....	111	Schlottmann, Robert S. ....	283
Geyer, L. H. ....	13	Sheehan, Nancy ....	75
Goldberg, Carlos ....	323	Sherman, Martin F. ....	317
Gough, Harrison G. ....	199	Shihadeh, Emile ....	335
Graham, William K. ....	55	Smeets, Paul M. ....	119
Gray, Roger T. ....	173	Striefel, Sebastian ....	119
Grodzitsky, Phyllis ....	189	Vincent, Clark E. ....	33
Gulas, Ivan ....	319	Von Wright, J. M. ....	9
Hadar, Ilana ....	105	Weiner, Michael Jay ....	195
Haney, C. Allen ....	33	Wuebben, Paul L. ....	145
		Zuckerman, Miron ....	227

# TABLE OF CONTENTS

Judgment of relative recency: Developmental trends . . . . .	3
By J. M. VON WRIGHT	
Manipulation of situational favorability in tests of the contingency model .	13
By L. H. GEYER AND J. W. JULIAN	
The effects of a zero interval on semantic differential rotated factor loadings	23
By JOHN H. FORTHMAN	
Motivational determinants of family planning clinic attendance . . . .	33
By CARL M. COCHRANE, CLARK E. VINCENT, C. ALLEN HANEY, AND ROBERT MICHIELUTTE	
Name style and conservatism . . . . .	45
By ROGER BOSHIER	
Polarity and "accuracy" of ratings and the meaningfulness of personality di- mensions . . . . .	55
By WALTER C. BORMAN AND WILLIAM K. GRAHAM	
Premackian reinforcement of classroom behavior through topic sequencing .	61
By JACK C. HARTJE	
Conservation of space in noninstitutionalized old people . . . . .	75
By DIANE E. PAPALIA, ELIZABETH KENNEDY, AND NANCY SHEEHAN	
Effects of external genital sensory feedback on copulatory behavior of rats .	81
By PETER O. PERETTI	
Pronunciability in verbal learning of the deaf . . . . .	89
By KATHLEEN CHEN	
Self-image disparity and attachment to ethnic subculture . . . . .	97
By MARK MCCORMICK AND DAVID BALLA	
Self-perception differences among kibbutz and city adults in Israel and Jewish and non-Jewish adults in the United States . . . . .	105
By LEO A. PIROJNIOFF AND ILANA HADAR	
Lysergic acid diethylamide (LSD 25): XXXX. Effect of pH on transport of methysergide and LSD 25 across gill membrane . . . . .	111
By H. H. GETTNER, P. A. CARONE, AND H. A. ABRAMSON	
The effect of experimenter absence and response delay on nonreinforced imitation . . . . .	119
By PAUL M. SWEETS AND SEBASTIAN STRIEBEL	
An investigation of the relationship between stimulus explicitness and entering behavior in facilitating student achievement . . . . .	129
By THOMAS C. ARNOLD AND FRANCIS M. DWYER	
An exploratory study of individual differences in perceptual centering and de- centering . . . . .	133
By ARIE COHEN AND FRANK H. FARLEY	

Creative thinking ability, school readiness, and intelligence in first grade children By NICHOLAS C. ALIOTTI AND WILLIAM E. BLANTON	137
A critique of Sarnoff and Zimbardo's psychoanalytic alternative to a social comparison theory of emotions . . . . .	145
By PAUL L. WUEBBEN	
Fertility ratio: Its relationship to mental ability, school achievement, and race By R. T. OSBORNE	159
Changes in college student attitudes toward the Arab-Israel, India-Pakistan, and Vietnam conflicts . . . . .	165
By WILLIAM B. ROSENBAUM AND LEONARD L. ROSENBAUM	
Change-making strategies in children and adults . . . . .	173
By HOWARD R. POLLIO AND ROGER T. GRAY	
Disparagement by a subordinate, ingratiation, and the use of power . . .	181
By EUGENE M. FODOR	
Sex-role identification and humor among preschool children . . . . .	189
By PAUL E. MCGHEE AND PHYLLIS GRODZITSKY	
The "lost letter technique" as a predictor of the 1972 presidential election .	195
By MICHAEL JAY WEINER AND EDWARD LUREY	
A factor analysis of contraceptive preferences . . . . .	199
By HARRISON G. GOUGH	
Parental dominance and sex-role identification in schizophrenia . . . .	211
By NICHOLAS D. KOKONIS	
Psychological test performance and indigenous conceptions of intelligence .	219
By ROBERT E. KLEIN, HOWARD E. FREEMAN, AND RICARDO MILLETT	
Signature size and self-estimation: A brief note . . . . .	223
By E. R. MAHONEY	
Status of frustrator as a facilitator of aggression: A brief note . . . .	225
By LAWRENCE J. CHASE AND NORBERT H. MILLS	
A political dimension on the I-E scale: A brief note . . . . .	227
By MIRON ZUCKERMAN	
The development of personal space schemata toward body build . . . .	229
By RICHARD M. LERNER	
Perception of environmental modifiability and involvement in antipollution activities . . . . .	237
By HANNA LEVENSON	
Developmental trends in wishes, confidence, and the sense of personal control from childhood to middle maturity . . . . .	241
By JOHN E. HORROCKS AND MILTON C. MUSSMAN	
The effect of first names on conflicted decisions: An experimental study . .	253
By THOMAS V. BUSSE AND CRAIG LOVE	
Language dominance and bilingual recall . . . . .	257
By STARRETTE DALTON	

Scaling meaningfulness (M) of trigrams with children and retardates . . . . .	267
BY CECIL O. CAMPBELL AND ALFRED A. BAUMEISTER	
Practice effect in anagram solving . . . . .	279
BY EDWARD I. GAVURIN	
The effects of verbal reports of violence on aggression . . . . .	283
BY WILLIAM H. SCHARFF AND ROBERT S. SCHLOTTMANN	
Studies in the neurophysiology of learning: VIII. Oscillatory potentials resulting from cerebral self-stimulation in rats . . . . .	291
BY J. A. GENDERELLI	
Locus of control and self-disclosure of public and private information by college men and women: A brief note . . . . .	317
BY RICHARD M. RYCKMAN, MARTIN F. SHERMAN, AND GARY D. BURGESS	
MMPI two-point codes for a "normal" college male population: A replication study . . . . .	319
BY IVAN GULAS	
Some effects of fear of failure in the academic setting . . . . .	323
BY CARLOS GOLDBERG	
Intellectual ability as related to age and stage of disease in muscular dystrophy: A brief note . . . . .	333
BY F. WILLIAM BLACK	
The perceptions of parolees and parole officers . . . . .	335
BY EMILE S. SHIHADDEH	
Concept-scale interaction with the semantic differential technique . . . . .	345
BY DAVID L. KLEMMACK AND JOHN A. BALLWEG	
Effects of IQ and mental age on verbal imitative performance of children . . . . .	353
BY REX FOREHAND, BRUCE ROBBINS, AND CHARLES PAT BRADY	





# ABBREVIATIONS OF WORDS USED IN THE TITLES OF JOURNALS

(One-word titles are never abbreviated. The word "the" is not used, nor its equivalent in any other language. The word "of" or its equivalent in other languages is used only to discriminate what would otherwise be identical titles in different languages. The word "and" is always used, but indicated by "&" in the Roman alphabet. Only English words are indicated here, but the corresponding words in other languages should receive a corresponding abbreviation. All abbreviations and all one-word titles should be in italics.)

Abnormal .....	<i>Abn.</i>	Japanese .....	<i>Jap.</i>
Abstracts .....	<i>Abst.</i>	Journal .....	<i>J.</i>
American .....	<i>Amer.</i>	Mathematical .....	<i>Math.</i>
Anatomy .....	<i>Anat.</i>	Measurement .....	<i>Meas.</i>
Animal .....	<i>Anim.</i>	Medical .....	<i>Med.</i>
Applied .....	<i>Appl.</i>	Mental .....	<i>Ment.</i>
Archives .....	<i>Arch.</i>	Monographs .....	<i>Monog.</i>
Association .....	<i>Assoc.</i>	Neurology .....	<i>Neurol.</i>
Attitude .....	<i>Attit.</i>	Opinion .....	<i>Opin.</i>
Australian .....	<i>Aust.</i>	Orthopsychiatry .....	<i>Orthopsychiat.</i>
Behavior .....	<i>Behav.</i>	Personality .....	<i>Personal.</i>
British .....	<i>Brit.</i>	Personnel .....	<i>Person.</i>
Bulletin .....	<i>Bull.</i>	Philosophy .....	<i>Philos.</i>
Bureau .....	<i>Bur.</i>	Physics .....	<i>Phys.</i>
Canadian .....	<i>Can.</i>	Physiology .....	<i>Physiol.</i>
Character .....	<i>Charac.</i>	Proceedings .....	<i>Proc.</i>
Children .....	<i>Child.</i>	Psychiatry .....	<i>Psychiat.</i>
Chinese .....	<i>Chin.</i>	Psychoanalysis .....	<i>Psychoanal.</i>
Clinical .....	<i>Clin.</i>	Psychology .....	<i>Psychol.</i>
College .....	<i>Coll.</i>	Psychosomatic .....	<i>Psychosomat.</i>
Comparative .....	<i>Comp.</i>	Quarterly .....	<i>Quart.</i>
Consulting .....	<i>Consult.</i>	Religious .....	<i>Relig.</i>
Contributions .....	<i>Contrib.</i>	Research .....	<i>Res.</i>
Development .....	<i>Devel.</i>	Review .....	<i>Rev.</i>
Educational .....	<i>Educ.</i>	School .....	<i>Sch.</i>
Experimental .....	<i>Exper.</i>	Science .....	<i>Sci.</i>
General .....	<i>Gen.</i>	Social .....	<i>Soc.</i>
Genetic .....	<i>Genet.</i>	Statistics .....	<i>Stat.</i>
Indian .....	<i>Ind.</i>	Studies .....	<i>Stud.</i>
Industrial .....	<i>Indus.</i>	Teacher .....	<i>Teach.</i>
International .....	<i>Internat.</i>	University .....	<i>Univ.</i>
Italian .....	<i>Ital.</i>		

# Preparation of Manuscripts for The Journal Press

## GENERAL INSTRUCTIONS

1. The proper sequence for the parts of your submitted manuscript is as follows: (a) text, (b) references, (c) footnotes, (d) tables, (e) figures, and (f) figure legends. However, monographs start with a table of contents and may have an acknowledgment page before the text and an appendix immediately after the text.
2. Use heavy typewriter paper,  $8\frac{1}{2} \times 11$  inches, double-space *all* lines, and leave margins for editorial work. Do not use onionskin, odd sizes, and abrasive or wax finishes.
3. Submit original typewritten version and one copy. Retain second copy for proofing.
4. Retype any page on which written corrections have been made.
5. Do not begin a sentence with a numeral.
6. A summary at the beginning of the text is required for articles over 500 words.
7. Each quotation should indicate the page number of the original source. The original publisher must give permission for lengthy quotations and use of tables or figures.
8. Do not fold your manuscript.
9. Enclose a submission letter, with a statement that the manuscript is not under consideration elsewhere. If you are unknown to the Editors, kindly give your credentials.

## FORMAT AND SPECIFIC INSTRUCTIONS

### A. TEXT DIVISIONS

#### I. THE TITLES OF JOURNAL ARTICLES AND THE MAJOR SUBDIVISIONS OF MONOGRAPHS ARE PRINTED IN TEN-POINT CAPS CENTERED ON THE PAGE

- A. THE NEXT SUBDIVISION TITLE IS PRINTED IN CAPS AND SMALL CAPS CENTERED ON THE PAGE
1. *Then italics, with Principal Words, Upper and Lower Case, Centered on the Page*
  - a. *Then italics, upper and lower case, 1-em run-in side head.*
    - (1). *Then italics, upper and lower case, 2-em run-in side head.*
    - (a). *Then italics, upper and lower case, 3-em run-in side head.*
- [Further subdivision should be merged into the text without marginal indentation, and should be numbered with small letters.]

### B. REFERENCES

References should be arranged in alphabetical order by author, numbered and referred to in the text by number (2). Double-space!

The proper form of a book reference is as follows:

2. DOE, J. The Preparation of Manuscripts. New York: Holt, 1963. Pp. 400-418.

The proper form of a journal reference is as follows:

2. DOE, J. The preparation of manuscripts. *J. Gen. Psychol.*, 1963, 68, 450-462.

If in the text it is desirable to refer to a page, thus (2, p. 45).

### C. FOOTNOTES

Use as few as possible and number consecutively in the text thus.<sup>1</sup>

- <sup>1</sup> Footnote (on the separate footnote page). Double-space!

### D. TABLES

Each table should be typed on a separate sheet and should be Arabic numbered (Table 2). Each column requires a heading. Vertical lines should be avoided. Mention tables consecutively in the text and indicate approximate insertion points.

### E. FIGURES

Figures should be submitted as glossy prints of the approximate size for final reproduction (ordinarily 4-4½ inches in width). Lettering and lines should be sharp and clean. Figure should be Arabic numbered (Figure 2) in the text, and have consecutive mention and approximate insertion points. Each figure requires a legend, but all legends should be submitted double spaced on a separate Figure Legends page.

339.00 per annum  
\$18.00 per volume  
Single Numbers \$9.00

BIMONTHLY  
Three volumes per year  
Immediate publication

November, 1973  
Volume 85, Second Half

Founded in 1935 by Carl Murchison

# The Journal of PSYCHOLOGY

The General Field of Psychology

POWELL MURCHISON  
*Managing Editor*

HELEN M. DAVIS  
*Associate Editor*

## Editorial Board

JOHN A. ANTOINETTI  
Life Insurance Agency  
Management Association

S. HOWARD BARTLEY  
Michigan State University

ARTHUR G. BILLS  
University of Cincinnati

E. A. BOTT  
University of Toronto

LEONARD CARMICHAEL  
National Geographic Society

IRVIN L. CHILD  
Yale University

HERBERT S. CONRAD  
U. S. Office of Education

JOHN F. DASHKILL  
Florida State University

LEONARD W. DOOR  
Yale University

If this space should be unstamped, this is the regular library edition.  
But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not  
be resold for a period of five years from date of publication.

FRANK A. GELDARD  
Princeton University

ERNEST R. HILGARD  
Stanford University

JOHN E. HORROCKS  
Ohio State University

DAVID HOTHERSALL  
Ohio State University

HEINRICH KLÜVER  
University of Chicago

NORMAN L. MUNN  
University of Adelaide,  
Australia

WARREN H. TEICHNER  
New Mexico State  
University

LEE EDWARD TRAVIS  
Fuller Theological  
Seminary

JOSEPH ZUBIN  
Psychiatric Institute,  
Columbia University

Copyright © 1973 by The Journal Press  
Provincetown, Massachusetts, U. S. A.

All rights reserved

Second-class postage paid at Provincetown, Mass.  
and additional mailing offices

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

# Microfiche Now Available

## Back Issues

While continuing to maintain a complete inventory of paper-bound back issues for all five journals—some reprinted and some original editions—The Journal Press is now offering microfiche (negative silverfiche) versions for all but the last two calendar years: i.e., Volume 1 through 1970 volumes are available immediately; 1971 volumes will first be available in January 1974; 1972 volumes, in January 1975, etc. The minimum order is for one volume. The following rates are in effect: Paper-bound—single numbers, \$9.00; microfiche, not available. Paper-bound—two-part volumes, \$18.00; microfiche, \$6.00. Paper-bound—three-part volumes, \$27.00; microfiche, \$9.00. Paper-bound—four-part and over-four-part volumes, \$36.00; microfiche, \$12.00. The only exceptions (special issues) are starred below.\*

VOLUMES	DATES	PRICE PER VOLUME		PRICE THROUGH 1970	
		PAPER	MICRO-FICHE	PAPER	MICRO-FICHE
<i>Journal of Genetic Psychology</i> (Pedagogical Seminary)					
1-117	1891-1970			\$2754.00	\$895.00
1-4	1891-April 1897	\$27.00	\$ 9.00		
5-37	July 1897-1930	36.00	12.00		
38*	1930	18.00	6.00		
39	1931	36.00	12.00		
40 and later	1932-	18.00	6.00		
<i>Genetic Psychology Monographs</i>					
1-82	1926-1970			1872.00	620.00
1-22	1926-1940	36.00	12.00		
23 and later	1941-	18.00	6.00		
<i>Journal of General Psychology</i>					
1-83	1928-1970			1566.00	520.00
1-3	1928-1930	36.00	12.00		
4*	1930	18.00	6.00		
5	1931	36.00	12.00		
6 and later	1932-	18.00	6.00		
<i>Journal of Social Psychology</i>					
1-82	1930-1970			1656.00	550.00
1-10	1930-1939	36.00	12.00		
11 and later	1940-	18.00	6.00		
<i>Journal of Psychology</i>					
1-76	1935/36-1970			1368.00	450.00
1 and later	1935/36-	18.00	6.00		

A mailing charge of \$0.30 per volume is made for foreign delivery of paper editions. Domestic paper mailings carry no postage charge. Prepaid foreign and domestic microfiche orders are also postage free; other orders will be billed for costs. Remittances should be by check or money order payable in U.S. funds on a United States bank, with no bank commission deducted. Advance payment is required from all individuals and from foreign purchasers. Microfilm versions and special bindings are not available.

Microfiche orders only may be sent directly to The Journal Press or to Johnson Associates Inc., 175 Fifth Avenue, New York, New York 10010.

THE JOURNAL PRESS

2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.



*Notice: All copyrighted material reprinted or otherwise reproduced from Journal Press publications should be given an appropriate acknowledgment. No permission is necessary for brief quotations and originally worded abstracts (less than 300-word total or 10 percent of length, whichever is least, for any one article). Written permission is required for reprinting or copying tables, figures, summaries, entire articles, or substantial portions thereof; such permission normally carries a charge of \$5 per page and is contingent upon agreement and possible stipulation of the original author or authors. Estate permission can be required. Authors requesting use of their own material are not charged. Fees may be waived for limited-copy instructional usage and for government periodicals. Special rates can be applied when justified.*

# The Journal of Psychology

Founded in 1935 by Carl Murchison

The Journal of Psychology was founded at the request of many psychologists throughout the world for the purpose of establishing a medium in which publication shall be immediate and in which other publication experiments may be carried out.

(1). The subscription price is \$39.00 per annum, \$18.00 per volume, \$9.00 per issue (three volumes per year).

(2). Manuscripts may be sent to any member of the Editorial Board or may be sent directly to The Journal Press. The lag in publication should not be more than a few weeks.

(3). The author will receive 200 reprints of his paper, and subscribers will receive the accumulated publications bound into half-volume units. The second half of each volume will contain the title page and table of contents for that volume. Each half-volume unit will be mailed to the subscribers at its stated date.

(4). Each author will be charged the manufacturing costs of any zincs, halftone cuts, tables, or equations. Excessive use of special fonts (references, footnotes, quotations, appendices) will also be charged.

(5). The publication charges are as follows:

1 to 4 pages .....	\$60.00
5 to 8 " .....	120.00
9 to 12 " .....	180.00
13 to 16 " .....	240.00
17 to 20 " .....	300.00
21 to 24 " .....	360.00
25 to 28 " .....	420.00
29 to 32 " .....	480.00

These charges are subject to adjustment as justified.

(6). The accumulated units will be mailed to subscribers bimonthly: January, March, May, July, September, and November. The units may vary in size to accommodate publication needs.

All back issues are available. See announcement on inside front cover.

THE JOURNAL PRESS

2 Commercial Street

Provincetown, Massachusetts 02657

U. S. A.

# Journal of Genetic Psychology

Founded by G. Stanley Hall in 1891; reorganized by Carl Murchison in 1925

Quarterly, beginning with March. Devoted to developmental and clinical psychology, with briefly reported Replications and Refinements and occasional book reviews. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$26.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

---

# Journal of General Psychology

Founded in 1927 by E. B. Titchener and Carl Murchison

Quarterly, beginning with January. Devoted to experimental, physiological, and comparative psychology, with briefly reported Replications, Refinements, and Comments. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$26.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

---

# Journal of Social Psychology

Founded in 1929 by John Dewey and Carl Murchison

Bimonthly, beginning with February. Devoted to studies of persons in group settings and of culture and personality; special attention to cross-cultural articles and notes, to field research, and to briefly reported Replications and Refinements. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Three volumes annually. Per annum, \$39.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

---

*Each author will receive 100 unbound reprints gratis. Additional reprints may be purchased at a nominal cost when the author returns page-proof.*

THE JOURNAL PRESS  
2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.

# Club Rate for Five International Psychological Journals

	PER ANNUM
THE JOURNAL OF GENETIC PSYCHOLOGY . . . . . (Quarterly, two volumes per year)	\$26.00
GENETIC PSYCHOLOGY MONOGRAPHS . . . . . (Quarterly, two volumes per year)	\$26.00
THE JOURNAL OF GENERAL PSYCHOLOGY . . . . . (Quarterly, two volumes per year)	\$26.00
THE JOURNAL OF SOCIAL PSYCHOLOGY . . . . . (Bimonthly, three volumes per year)	\$39.00
THE JOURNAL OF PSYCHOLOGY . . . . . (Bimonthly, three volumes per year)	\$39.00
	<hr/>
	TOTAL \$156.00
TEACHERS AND STUDENTS EDITION, <i>Restricted Sale</i>	\$50.00

The *Teachers and Students Edition* is open to any individual psychologist, teacher, student, or physician but requires an agreement that the edition will not be offered for sale till after five years from publication and an agreement that the edition will not be substituted for an institutional subscription. Write for our Teachers and Students contract card.

THE JOURNAL PRESS  
2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.

## Genetic Psychology Monographs

Founded in 1925 by Carl Murchison

Quarterly, beginning with February. Devoted to developmental and clinical psychology. Each number one or more complete researches. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$26.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

**NOTICE.** Effective immediately we offer prior publication for certain selected monographs. Charges will be based on the announced rates for *The Journal of Psychology*. Regular monographs will carry no page charge for the first 50 pages, but for all pages above 50 there will be a charge of \$10.00 per page. Manufacturing costs of any zincs, halftone cuts, tables, or equations will be charged to authors, as will excessive use of special fonts. Two hundred reprints of each monograph will be supplied free of charge.

The listing of monographs for 1926 through 1963 will be printed in the first half of each volume of all five journals. Monographs published after 1963 will be listed in the second half of each volume.

# Genetic Psychology Monographs (continued)

## VOLUME 69—January-June, 1964

- Occupational and industrial therapy in the treatment of psychiatric patients: A controlled study of efficacy in an intensive-treatment institution—M. HUNDZIAK AND B. PASAMANICK  
An analysis of Carl R. Rogers' theory of personality—M. S. KRAUSE  
Personality aspects of teaching: A predictive study—H. ZIMILES, B. BIBER, W. RABINOWITZ, AND L. HAY  
Situational and personal antecedents of incipient alienation: An exploratory study—W. H. JARRETT AND A. O. HALLER
- Age changes in children's Mosaic responses from five to ten years—L. B. AMES AND F. L. ILG  
A study of the interview process: The caseworker-client relationship—H. J. MEYER, E. F. BORGATTA, AND D. FANSHILL  
A study of introversion-extroversion as a personality dimension—R. M. ALLEN, H. M. RICHNER, AND R. J. PLOTNIK  
Psychological testing with Northwest Coast Alaskan Eskimos—C. E. PRESTON  
A study in family pathology—M. B. HANS

## VOLUME 70—July-December, 1964

- Structure and function—H. GUTMAN  
The Lowenfeld Mosaic Test: Norms for five-to-ten-year-old American public-school children and comparative study of three groups—L. B. AMES, F. L. ILG, AND J. AUGUST  
Social and personality correlates of children's estimates of height—J. P. SHAPPER  
The development of ethnic attitudes in New Zealand school children—G. M. VAUGHAN
- The antecedents of two varieties of vocational interests—L. E. TYLER  
Personality in old age—P. E. SLATER AND H. A. SCARR  
Variables affecting experimenter bias in a group situation—R. ROSENTHAL, C. J. FRIEDMAN, C. A. JOHNSON, K. FODE, T. SCHILL, R. C. WHITE, AND L. L. VIKAN  
Ordinal position and sex role identification—B. G. ROSENBERG AND B. SUTTON-SMITH  
Personality correlates of accuracy of self-appraisal in four-year-old children—PAUL S. WEINER

## VOLUME 71—January-June, 1965

- Author index for Genetic Psychology Monographs 1926-1964—L. W. FERGUSON  
A theoretical, developmental, and clinical investigation of embarrassment—J. M. SATTLER  
School readiness as evaluated by Gesell developmental, visual, and projective tests—F. L. ILG, L. B. AMES, AND R. J. APEL  
Rorschach scales of regression in psychosis—S. NEICKER, A. G. SLEMON, AND D. A. QUIRK  
The effects of extra tactile stimulation on a group of institutionalized infants—L. CASLER  
Human-cyloid conditioning: The recent experimental literature—L. E. ROSS AND T. F. HARTMAN
- Social status and schoolroom aggression in third-grade children—R. TOIGO  
Overt and covert homosexual problems among male college students—L. J. BRAATEN AND C. D. DARLING  
Role conceptions of the psychiatric aid—M. P. LAWTON AND A. E. GOLDBMAN  
Changes in study of values scores during the four years of college—C. W. HUNTLEY

## VOLUME 72—July-December, 1965

- A study of process and method in three-year-old twins and triplets learning to read—W. FOWLER  
Individualized teaching with therapeutic aims: A methodological study—M. KOHN AND M. RUDNICK  
Ability and social adjustment at midlife of persons earlier judged mentally deficient—E. L. MILLER
- Development of incentive values in childhood—S. L. WITRYOL, D. J. TYRRELL, AND L. M. LOWDEN  
Longitudinal changes in intolerance and authoritarianism for subjects differing in amount of college education over four years—W. T. PLANT  
An NDEA counseling and guidance institute: Prediction, performance, and follow-up—R. H. DANA  
The influence of language on classification: A theoretical model applied to normal, retarded, and deaf children—H. C. FURTH AND N. A. MCGRAM  
Time estimates as affected by incentive class and motivational level—M. FELIX  
Clinical psychology: An historical survey—V. S. SEXTON

## VOLUME 73—January-June, 1966

- Vulnerability in infants and young children: A pilot study—G. HEIDER
- Racial preference and ego-identity among white and Bantu children in the Republic of South Africa—A. J. GREGOR AND D. A. MCPHERSON  
Mother-child interaction in cerebral palsy: Environmental and psychosocial obstacles to cognitive development—E. SHERR AND R. KASTENBAUM  
Children's stories—L. B. AMES

## VOLUME 74—July-December, 1966

- Changes in personality for groups completing different amounts of college over two years—W. T. PLANT AND C. W. TELFORD  
Brain dynamism in perception of reversible figures—G. K. YACORZYNSKI  
Changes in Rorschach response throughout the human life span—L. B. AMES  
Mnemonic virtuosity: A study of chess players—A. BINET (Translated by M. L. SIMMEL AND S. B. BARRON)
- The development of behavior patterns in children—B. B. TYLER, F. B. TYLER AND J. E. RAFFERTY  
Correlates of classroom functioning—L. M. SMITH AND B. B. HUDGINS  
Children's behavior syndromes and parental responses—R. L. JENKINS, E. NUREDDIN, AND I. SHAPIRO

## VOLUME 75—January-June, 1967

- Studies of an elementary game of strategy—B. SUTTON-SMITH AND J. M. ROBERTS  
Child care practices anteceding three patterns of preschool behavior—D. BAUMRIND  
The impact of cognitive maturity on the development of sex-role attitudes in the years 4 to 8—L. KOHLBERG AND E. ZIGLER
- Differential factor structures in mathematical ability—P. S. VEKY  
Perception of the vertical as a function of affective stimuli and extraneous stimulation in schizophrenic and normal subjects—J. O. WELLS AND W. E. CALDWELL  
Mid-life attainment of the mentally retarded: A longitudinal study—W. R. BALLER, D. C. CHARLES, AND E. L. MILLER

## VOLUME 76—July-December, 1967

- Scalogram analysis of some developmental sequences of infant behavior as measured by the Bayley Infant Scale of Mental Development—R. KOHN-RAZ  
The leader and the lost: A case study of indigenous leadership in a poverty program community action committee—L. A. ZURCHER  
A developmental study of the perception of people—R. N. WOLMAN  
Human fears: A factor analytic study of three age levels—G. W. RUSSELL
- Correlates of voluntary social participation—J. E. TEELER  
Personality characteristics and developmental history of creative college women—R. HESLON  
A systematic study of therapeutic interventions in nursing homes—B. A. STOTSKY

## VOLUME 77—January-June, 1968

- The fulfillment of promise: 40-year follow-up of the Terman Gifted Group—M. H. ODEN  
A scale for parents' ratings: Some ipsative and normative correlations—R. N. WALKER  
A methodological investigation of Piaget's concept of conservation of substance—A. C. BITTNER AND M. M. SHENKLE
- Sex differences in mental and behavioral traits—J. E. GARAI AND A. SCHREINFELD



# Genetic Psychology Monographs (continued)

## VOLUME 78—July-December, 1968

1. A theoretical and developmental investigation of delayed speech feedback—J. E. WATERS  
Growth changes in factorial structure by age and sex—N. W. DYE AND P. S. VERNY  
A multispect investigation of masculine development in kindergarten age boys—H. B. BILLER
2. The optimum limits of testing on the Wechsler Intelligence Scales—M. Y. QUERREHI  
A developmental study of the ego-ideal—L. VAN DEN DAKKE  
The relationship between physical condition immediately after birth and mental and motor performance at age four—N. EDWARDS

## VOLUME 79—January-June, 1969

1. Adolescents: From the views of the members of an informal adolescent group—T. P. JOSEPH  
Acculturation and childhood accidents among Chinese and Japanese Americans—M. KUROKAWA
2. Dimensions of achievement-related behavior among lower-class Negro parents—D. SOLOMON, R. J. PARELIUS, AND T. V. BUSSE  
Dimensions in the judgment of illness—J. S. KERRICK  
Reversal shifts and problem-solving abilities at ages 4-7—S. B. KASTNER  
The judgment of body landmark heights—H. NASH  
Recognition of body surface regions—H. NASH

## VOLUME 80—July-December, 1969

1. Animal aggression: Review of research—W. M. VERNON  
The effects of positive and negative reinforcement and test anxiety on the reading performance of male elementary school children—S. COTLER  
Personality and psychosomatic illness: A selective review of research—P. R. ROBBINS  
Task and motivational influences on the assessment of creative and intellectual ability in children—N. KOGAN AND F. T. MORGAN  
A study of an interview method for identifying family authority patterns—R. W. CHAMBERLIN
2. Conditioned reinforcing and aversive aspects of the stimuli defining the components of a two-component chain—A. KAUFMAN AND A. BARON  
The Henry Htleston Center Family Interaction Scales—M. L. BEHRENS, D. I. MEYERS, W. GOLDFARB, N. GOLDFARB, AND N. D. FIELDSTEEL

## VOLUME 81—January-June, 1970

1. Attitudes of college students and their changes: A 37-year study—V. JONES  
Tattooing among male and female offenders of different ages in different types of institutions—A. J. W. TAYLOR
2. Sex differences in mental health—J. E. GARAI  
A selective review of dominance-subordination relations in animals—D. VAN KREVELD  
The effects of visual or auditory stimulation given in infancy upon subsequent preference for stimulation of both modalities—W. R. BERKOWITZ  
The influence of certain child-rearing factors upon the development of a structured and salient sense of time—D. ZERN

## VOLUME 82—July-December, 1970

1. Child care attitudes and emotional disturbance among mothers of young children—B. J. COHLEN, J. L. WEISS, AND H. U. GRUNESBAUM  
A study of intragroup and intergroup relations among experimentally produced small groups—L. N. DIAB  
The development of role knowledge—D. R. HEISE AND E. P. M. ROBERTS
2. Middlecence: Age related stress periods during adult years—J. E. HORROCKS AND M. C. MUSSMAN  
The first six years of life: A report of current research and educational practice—E. LACROSSE, P. C. LEE (Ed.), F. LITMAN, D. M. OGILVIE, S. S. STODOLSKY, AND B. L. WHITE

## VOLUME 83—January-June, 1971

1. Measurement of parents' perceptions of their children's development—W. WOLVENBERGER AND R. A. KURTZ  
The concomitant development of cognitive and moral modes of thought: A test of selected deductions from Piaget's theory—L. C. LEE  
Children's influences upon parental behavior: An attempt to define the relationship with the use of laboratory tasks—J. D. OSORSKY
2. Parent behavior and child academic achievement, achievement striving, and related personality characteristics—D. SOLOMON, K. A. HOULIHAN, T. V. BUSSE, AND R. J. PARELIUS  
Memory factors and school success of average and special groups of ninth-grade boys—M. MEEKER AND C. E. MEYERS  
Personality correlates of vocational planning—R. BARNETT

## VOLUME 84—July-December, 1971

1. Perceptual distortion as a function of stimulus objects, sex, naivete, and trials using a portable model of the Ames Distorted Room—T. A. GERACE AND W. E. CALDWELL  
Galvanotactic behavior and reinforcement of fish *Brachydanio Rerio*—W. F. FLANIGAN, JR., AND W. E. CALDWELL  
Self-actualization, types of personality, and neuroses and psychoses—R. DREISTADT  
Children's awareness of their bodily appearance: A developmental study of factors associated with the body percept—E. GELLERT, J. S. GIRCUS, AND J. COHEN
2. Biogenetic phenomena: Some implications for the behavioral and neurophysiological sciences—F. RUSSO AND W. E. CALDWELL  
Sociometric task leadership and the perception of deviance and subsequent conformity as functions of personality—W. L. LIBBY, JR.  
Murderers who plead insanity: A descriptive factor analytic study of personality, social, and history variables—M. W. KAHN

## VOLUME 85—January-June, 1972

1. Children's understanding of negation as a logical operation—S. S. FELDMAN  
Selective auditory perception of parental voices by the disturbed child—S. SANDERS  
Identity type, sociometric ratings, and group behavior—C. N. ZIMET AND C. SCHNEIDER
2. Patterns of perceived reaction to frustration among adolescents—J. O. PALMER  
The need for influence: Development and validation of a measure, and comparison with the need for power—J. S. ULEMAN  
C. Lloyd Morgan, E. C. Tolman, and comparative method—E. NEWBURY  
Relearning motor skills in hemiplegia: Error analysis—L. DILLER, J. BUXBAUM, AND S. CHIOTELIS  
Effects of participation in rule formation on the moral judgment of children—R. L. MERCHANT AND F. REBELSKY  
Discrepancy and attention in the five-month infant—C. M. SUPER, J. KAGAN, F. J. MORRISON, M. M. HAITH, AND J. WEIPFENBACH



# Genetic Psychology Monographs (continued)

## VOLUME 86—July-December, 1972

1. The role of enduring orientations to the environment in personality development—W. C. BRONSON  
Is there an error of the standard? A critique of Piaget's perceptual theory—D. A. BEGELMAN AND G. J. STEINFELD  
A cross-cultural study of the relationship between resistance to temptation and maternal communication—E. R. GREENGLASS  
The intellectual and achievement effects of preschool cognitive stimulation of poverty Mexican-American children—W. T. PLANT AND M. L. SOUTHERN
2. Factors affecting agnostic behavior in several subhuman species—R. F. GOTTIER  
Differential behavioral characteristics of nonpromoted children—J. A. CHASE  
An analysis of cognitive-developmental sequences—J. H. FLAVELL

## VOLUME 87—January-June, 1973

1. A survey and interpretation of article characteristics: *The Journal of Genetic Psychology* 1945-1969—J. E. HORROCKS AND J. D. HOGAN  
The development of listener adapted communications in grade-school children from different social-class backgrounds—K. T. ALVY  
Locus of control among schizophrenics and other hospitalized psychiatric patients—T. F. CASH AND J. J. STACK  
Interactive styles and social adaptation—C. N. EDWARDS
2. The motivation of exploration strategies in young children—J. D. COLE  
Children's representation of the human figure: The effects of models, media, and instructions—C. GOLOMB  
A perceptual program for classroom teachers: Some results—R. N. WALKER AND J. STREFF  
Behavioral approaches to study and treatment of psychogenic tics—M. HERSEN AND R. M. EISLER  
Patterns of intellectual development during childhood and adolescence—M. Y. QUERESHI

## VOLUME 88—July-December, 1973

1. Adolescent marginality—A further study—J. H. BAMBER  
Psychological problems of abortion for the unwed teenage girl—C. D. MARTIN  
Perceptual preferences in relation to aspects of personality—L. M. GOODMAN  
An exploratory study of personality manifestations in paintings—R. B. BURT

THE JOURNAL PRESS  
2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.



## PUBLICATION IN THE JOURNAL OF PSYCHOLOGY

*The Journal of Psychology* provides a special service to authors and readers. Articles accepted for publication appear in print with the least possible delay. Although editorial standards are the same as those for the other journals, all correspondence and handling are expedited, and the referees may, at their discretion, permit some latitude in length of treatment with respect to tabular material, discussion, and review of the literature.

To save time, authors are requested to exercise unusual care in preparing their papers. Special attention should be paid to the General Instructions on the back cover of each issue and to the announcement on the first preliminary page.

\$39.00 per annum  
\$18.00 per volume  
Single Numbers \$9.00

BIMONTHLY  
Three volumes per year  
Immediate publication

November, 1973  
Volume 85, Second Half

(Founded in 1935 by Carl Murchison)

# The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a bilateral contract, and may not be resold for a period of five years from the date of publication.

NOVEMBER, 1973

(Manuscripts are printed in the order of final acceptance)

- Mental health and population density . . . . . 171  
By J. T. LAIRD
- Measuring social relationships in emotionally disturbed boys . . . . . 179  
By CHARLES SCHAEFER
- Time judgments by magnitude estimation and magnitude production and anxiety: A problem of comparison between normals and certain schizophrenic patients . . . . . 187  
By JACQUES RUTSCHMANN
- A preliminary note on demographic and personality correlates of decubitus ulcer incidence . . . . . 225  
By JOHN G. CULL AND OTHO H. SMITH

(OVER)

Copyright © 1973 by The Journal Press  
Provincetown, Massachusetts, U. S. A.

All rights reserved

Second-class postage paid at Provincetown, Mass  
and additional mailing offices

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

The relation of student and teacher traits of authoritarianism to student achievement in English . . . . .	229
By JOHN G. WAY	
The structure of behavioral values of college students . . . . .	235
By JOHN PAUL MCKINNEY	
The impact of exam question order effects on student evaluations . . . . .	245
By MICHAEL L. DEAN	
Personality and attitudes: A re-emphasis upon the cognitive component . . . . .	249
By SALLY M. DUNLAP, SAMUEL L. GAERTNER, AND A. DAVID MANGELSDORFF	
Components of depression in attempted suicides . . . . .	257
By AARON T. BECK AND DAVID LESTER	
Conceptual systems and philosophical orientation . . . . .	261
By JOHN D. WILLIAMS AND JOHN H. KELLEHER	
Episodic analysis of novels . . . . .	267
By ROBERT R. SEARS AND DEBORAH LAPIDUS	
Lexical marking effects in the semantic differential . . . . .	277
By ANDREW R. GILPIN	
The stereotypes of four ethnic groups . . . . .	287
By MENI KOSLOWSKY	
The relative effectiveness of two methods of presenting visualized instruction . . . . .	297
By FRANCIS M. DWYER	
Massed <i>versus</i> spaced sessions in systematic desensitization . . . . .	301
By JOHN O. KLING AND ELLEN M. MINOGUE	
Managerial attitudes toward hiring ex-convicts . . . . .	305
By ALBERT N. B. NEDD	
Effects of interitem associative strength, rehearsal, and proactive inhibition on the retention of paired-associate lists . . . . .	313
By JOHN A. MILLS	
An examination of verbal imitative performance in young children . . . . .	323
By REX FOREHAND AND HAROLD L. GARDNER	
Frustration response categories and level of hostile expression . . . . .	329
By MORTON GOLDMAN AND CARYLN S. RHOADS	
Risk taking by individuals and informal groups with the use of industrial product purchasing situations as stimuli . . . . .	339
By PETER H. REINGEN	



## MENTAL HEALTH AND POPULATION DENSITY\*

*Saint John's University*

---

J. T. LAIRD

---

### SUMMARY

A comparison of mental illness and narcotic use in rural and urban areas indicates a greater incidence in areas with high population density. A hypothesis explored is that prior experiences condition some to be sanguine and others melancholic. Whereas both are vulnerable to stress, the reactions to signals indicating punishment will vary according to the conditioned temperament. The great frequency of signals in high population density areas activates a general stress reaction. Relief can be obtained for the melancholic by withdrawal, whereas for the sanguine it is by the use of drugs.

---

### A. INTRODUCTION

The compression of people within a defined territory may be altering significantly their inner milieu manifested by personal disruption of life style frequently referred to as alienation (literally, not to be in harmony with oneself—i.e., neurotic or character disorder; or, figuratively, to be a stranger to oneself—i.e., psychotic). Some symptoms expressed are feelings of hopelessness, despair, anxiety, futility, and depression. The diagnosis of an individual who has these persistent symptoms in a mild form is neurasthenia, and when extreme he is classified as having an incapacitating psychosis. (It was probably not by accident that psychiatrists were referred to as alienists, as these were symptoms encountered frequently.) In psychosis, the individual's level of consciousness is so fragmented that he can no longer maintain social goal orientated behavior, whereas in neurasthenia his level of consciousness is better integrated but imbued with anxiety while in the quest of goals. A third condition occurs when the individual attempts to mediate the symptoms contrary to socially approved methods, and then he is usually classified as having a behavioral or character disorder.

---

\* Received in the Editorial Office on July 12, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

For purposes of this study, individuals were grouped on the basis of their symptoms under two major categories: psychoses and behavioral disorders. Individuals whose social goal orientated and self-sustaining behavior was severely and markedly impaired because of mental illness were placed in the category of psychoses. These individuals were recommended for treatment in hopes of restoration of social functioning. Individuals who compulsively alter their level of consciousness by ingesting narcotic drugs, presumably as a self-induced homeostatic device, were placed in the category of behavioral disorders. Narcotic law violators were selected because they satisfy two essential requirements: (a) their behavior is persistently goal directed in spite of severe punitive consequences; (b) the incidence on a geographical basis for over the past 30 years is available. Individuals who use other drugs and alcohol, exhibit aggressive behavior, or manifest the mild symptoms of alienation were not considered because they often go officially undetected, and the syndromes, if capable of exact definition, have such elastic boundaries that categorization would be of questionable accuracy.

The hypothesis is that psychoses and behavioral disorders vary proportionally to population density.

## B. DEMOGRAPHIC STUDY I

### 1. *Rural Sample*

The most recent and extensive study of mental illness in a rural area was conducted in 1968 in Minnesota (6). The central quarter of the state, which extends from the border of South Dakota to Wisconsin, consists of 18,000 square miles with a population of 319,723. It is a rural area with an economic base of agriculture and small industries. The inhabitants are primarily of Scandinavian and German descent. There are no cities with a population in excess of 40,000. The population density is 17 inhabitants per square mile, whereas the population density of the U.S. is 57 per square mile.

A caseworker was assigned to each of 19 counties to review all records of the probate and district courts, welfare departments, mental health centers, and state hospitals to ascertain how many residents had been legally or medically diagnosed as mentally ill. This does not necessarily mean they were hospitalized, or even treated, but that they were diagnosed as being in need of psychiatric treatment in order to continue functioning or for restoration of functioning. In other words, they had a major mental illness best described as serious or incapacitating. As there are no private psychiatric hospitals or private psychiatrists in this area (there is one psycho-

analyst), it was considered that the incidence was close to the actual or identifiable rate of mental illness.

The second step was to have each caseworker in consultation with other professionals in the community estimate how many residents were seriously mentally ill but may not necessarily have been professionally diagnosed. Included in this group would be the ambulatory psychotic who receives private medical care or financial assistance. The incidence was the estimated mental illness of serious or incapacitating proportions. There were no age restrictions for either group, but those below 18 were categorized as children and the remainder as adult.

## 2. Urban Sample

The urban sample selected was the extensive Midtown Manhattan Study conducted by Srole *et al.* (11) because the category of mental illness as severe or incapacitating was available. Also, the population density of 25,000 per square mile in keeping with the 15,000 to 25,000 of Newark, Baltimore, Chicago, etc. The age range was restricted to those from 20 to 59.

## 3. Results and Discussion

The Minnesota study indicated that there were a total of 1095 individuals professionally diagnosed as severely mentally ill or about three-tenths of one percent of the population. In terms of estimated severe mental illness, there were a total of 4366 or slightly more than one percent of the population. In contrast, the Midtown Manhattan Study tabulated the incidence as slightly over 10 percent or 10 times higher than the rural figures despite the fact that the rural included children and those over 60, which the urban sample excluded.

The question is twofold: Are these findings an artifact? Is mental illness proportional to population density?

In terms of proportionality, the best statistics are those prior to the early 1950's because up to that time state hospitals were the exclusive receiving stations for the mentally ill. Since then there has been a proliferation of other agencies coping with the mentally ill, a change in philosophy in which the psychotic is to be treated in the community, and the use of ataractic drugs, all of which tend to conceal the true incidences and make reliable tabulations difficult to come by unless naturalistic methods are used, such as the Midtown Manhattan and Minnesota studies.

Malzburg (8) examined first admissions to New York State public mental hospitals from September 1949 to September 1951. The average annual rate

for first admissions was 184 per 100,000 population for urban areas and 97.7 for rural areas. He stated that there was a general progressive increase of rates of first admissions from rural to small urban groups to large urban areas. He cites further evidence in which the admissions for schizophrenia were highest from New York City, next highest from middle urban cities, such as Buffalo and Rochester, and lowest from rural areas. It might be argued that purposeful sampling is not a valid technique. However, Rose and Steele (10) made a survey of all available studies of the incidence of mental illness from throughout the United States, and they found that in every instance there was a greater prevalence in urban areas and concluded *the larger the city, the higher the rate*.

The other question to be explored is whether the studies are artifacts. There are four hypotheses that claim that all such studies have built-in sampling errors, and each of these should be answered.

The first hypothesis states that those with mental illness fit unobtrusively into rural life and are not likely to be detected. If this is correct, then one would expect that patients from rural areas admitted to hospitals would have a longer period of illness as opposed to those from urban areas. The findings do not support this but indicate that those from rural areas are as, or more, quickly admitted as urban patients. Malzburg (8) examined first admissions to hospitals in New York for those diagnosed as schizophrenic, and 13.6% from the rural area had symptomatic histories five or more years prior to hospitalization, whereas the urban admissions had a corresponding percentage but with 12 or more years of prior symptomatic histories.

The second hypothesis concerns urban drift, which states that alienated individuals find life intolerable in rural areas, so they drift to the city where they are detected and hospitalized. What is not mentioned is that it is just as reasonable to assume a rural drift. However, Hollingshead and Redlick (4) have shown that urban drift is of doubtful validity. Class I and II families, which have a low rate of mental illness, are more likely to be mobile. Of those who were mentally ill, only 44% were native to the city. Class V families, which make up the bulk of the mentally ill, are less likely to be transient. Of those who were mentally ill, 61% were life-long residents of the city. If drift were responsible, just the opposite would be expected.

The third hypothesis is downward drift, which claims that the mentally ill drift downward socially into congested areas. However, Hollingshead found that only 1.3% of the mentally ill were in a lower social class than their family's origination.

The final hypothesis to be challenged is that the higher percentage of



foreign born or immigrants who inhabit cities suffer from cultural shock and, hence, become mentally ill. Malzburg and Hollingshead, in independent investigations, could find no relationship between nativity and presence or absence of schizophrenia or other mental illness.

### C. DEMOGRAPHIC STUDY II

One of the most extensive analyses of drug addiction in the United States was conducted by Lindesmith (7) in which he documents the generally accepted belief that drug addiction is highest in large cities or high population density areas. The use of statistics concerning drug use poses many problems. Federal narcotic law violations, while complete, are too gross. Non-federal narcotic law violations, which include every geographical area as well as incidence according to size of community, include marijuana violations which confuse the interpretation. However, prior to the great popularity of marijuana which began by the mid-1960's in the United States and Canada, most violations were for the so-called "hard drugs" (opiates and cocaine); e.g., in 1960 there were only 169 Federal marijuana violations recorded.

The *Uniform Crime Reports* cited by Lindesmith (7), of persons charged with nonfederal narcotic law violation in the United States, are most complete geographically for the year 1961 and 1962 but less extensive prior to these years. The pattern, however, since 1932 has remained the same; the larger the community, the greater the incidence, similar to the pattern in mental illness. Illustrative of this is the incidence in 1962 of violations per 100,000 inhabitants. For communities under 10,000 the incidence was 2.8; from 10,000 to 25,000 the rate was 5.1; from 25,000 to 50,000 the rate was 8.3; from 50,000 to 100,000 the rate was 13.5; from 100,000 to 250,000 the rate was 15; over 250,000 the rate was 61.7. It should be noted that the rate quadruples when the population exceeds 250,000. This population density reaction also holds for geographic regions where the rate quadruples in comparing New England to the Middle Atlantic States (New York, New Jersey, Pennsylvania) and increases tenfold comparing West North Central States (Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, and South Dakota) to the Middle Atlantic.

### D. CONCLUSION

Demographic data indicate that severe mental illness and narcotic use have a greater incidence in high population density areas and may even be proportional to the density. It could be suggested that this holds true for other



psychological syndromes, but categorization and measurement are less precise.

A rationale for the empirical data must rely upon hypothetical constructs. Infrahuman naturalistic studies by Carrigher (1), Christian [cited by Hall (3)], Errington [cited by Hall (3)], and Funkenstein (2) are convincing that a nonspecific stressor of density activates a general adaptation syndrome which simulates human mental aberrations. If such is the case for humans, then the question arises as to why some are prone to psychoses and others to behavioral disorders (on the assumption that genetic factors are evenly distributed in the population).

The phenotypical behavior of those who are psychotic differs from those who have behavior disorders. The former are characterized by a lack of initiative, or motivation, and in terms of the general adaptation syndrome are in a stage of exhaustion. They withdraw from social contact in its many varieties. The latter are characterized by active behavior, often socially disapproved, as if in a stage of resistance. A further observable distinction is that psychotics rarely use drugs—in fact, it is even difficult to get them to take medications—whereas those with behavioral disorders are prone to drug use. Hence, one group is basically passive and avoids drugs, whereas the other group is basically active and is drug prone.

Psychological studies have long classified temperaments as sanguine or melancholic, extravert or introvert, high dominant or low dominant. While temperament has generally been considered a genotype, Pavlov did state that not everyone fits into these categories, and conditioning could influence behavioral predispositions (9). Virzhikovsky and Mayorov [cited by Kimble (5)] demonstrated that early social experiences could lead to behavioral constellations similar to the sanguine or melancholic. If, then, active and passive behaviors are influenced by conditioning, it could be that such idiosyncratic behavior is in response to signals that indicate forthcoming punishment, and the individual responds with an emotion of fear or anger. The melancholic, through prior experiences of frequent punishment, is passive and withdraws as an attempt to avoid the fear aroused by incoming signals. The sanguine, who has experienced inconsistent consequences of incoming signals, reacts with anger to diminish the source of the threat. This could account for Pavlov's finding that the melancholic when confronted with many fear signals becomes immobile, and treatment of choice is rest and quiet. The sanguine when confronted by fear arousing signals is activated (anger), and here the treatment of choice is a calming drug (bro-

mide). In the former case, drugs would only further immobilize, but in the latter they reduce the anger.

In terms of population density, individuals in high density areas are exposed to more stimuli than those in low density areas. The increase in stimuli raises the proportion of fear signals, and on the basis of prior conditioning one type of temperament would find withdrawal negatively reinforcing, whereas for the other it would be drugs.

Psychoses and drug addiction as consequences of fear and anger reactions are together but one consideration of a complex constellation. Other contributing factors, such as genetics, social modeling, and patterns of class behavior, should also be kept in mind.

#### REFERENCES

1. CARRIGHER, S. Wild Heritage. Boston: Houghton Mifflin, 1965. Pp. 173-180.
2. FUNKENSTEIN, D. H. Physiology of fear and anger. *Sci. Amer.*, 1955, 192(20), 74-80.
3. HALL, E. Territorial needs and limits. *Nat. Hist.*, 1965 (Dec.), 74, 16.
4. HOLLINGSHEAD, A., & REDLICK, F. Social Class and Mental Illness: A Community Study. New York: Wiley, 1958. Pp. 245-246.
5. KIMBLE, G. A. Foundations of Conditioning and Learning. New York: Appleton-Century-Crofts, 1967. P. 41.
6. LAIRD, J. T. Mental illness and treatment in rural U. S. A. In A. B. Tulipan & D. W. Heyder (Eds.), *Outpatient Psychiatry in the 1970's*. New York: Brunner/Mazel, 1970. Pp. 49-60.
7. LINDESMITH, A. R. The Addict and the Law. New York: Vintage, 1967. Pp. 108-113.
8. MALZBURG, B. The distribution of mental diseases in New York State, 1949-1951. *Psychiat. Quart. (Suppl.)*, 1955, Part 2, 235.
9. PAVLOV, I. P. Conditioned Reflexes. New York: Dover, 1960. Pp. 285-300.
10. ROSE, A. M., & STEELE, H. R. Mental Health and Mental Disorder. New York: Norton, 1955.
11. SROLE, L., et al. Mental Health in the Metropolis. New York: McGraw-Hill, 1962.

St. John's University  
Collegeville, Minnesota 56321



## MEASURING SOCIAL RELATIONSHIPS IN EMOTIONALLY DISTURBED BOYS\*

*The Children's Village*

---

CHARLES SCHAEFER

---

### SUMMARY

The purpose of this study was to investigate the relationship between measures of social attraction and social protection. Social attraction was measured by the Kuethe felt figures technique, while social protection was measured by the *Shipwreck Test*. The results revealed that scores of emotionally disturbed children on the two measures were highly correlated. Another noteworthy finding was that the Ss showed more social attraction toward the grandfather, dog, and friend figures, than to the mother, father, and grandmother figures.

---

### A. INTRODUCTION

A major concern of social psychological theory (1) is self-other orientation: i.e., a person's perceived relationships with significant other people, such as parents, siblings, and friends. Apart from self-reports, the method for measuring these perceived social relationships often involves the arrangements of symbols or objects representing the self and significant other people. For example, one might use the physical distance between the subject's arrangements of paper cutout human figures as a measure of the social orientation of the subject; i.e., the closer the physical arrangement between self and other figures, the closer the social attachment of the two figures [see Kuethe (2, 3)]. Using Kuethe's techniques, Weinstein (7) found that emotionally disturbed children differed predictably from normal children. When asked to place pairs of cutout figures on flannel boards, normals placed child figures closer to mother than to father or peer figures; emotionally disturbed children did the reverse. An alternate technique for measuring social interest or concern is the *Shipwreck Test* (4, 5). On this test, the child is asked to pretend that he is on a sinking ship with five members of his family, together with his friend, dog, and television set. He is then

---

\* Received in the Editorial Office on July 17, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

asked which of these nine figures he would like to save by assigning them to one of the six available lifeboats. Thus, the *Shipwreck Test* is designed to measure "social protection" or concern for the safety of others, while the physical distance between cutout figures seems to measure "social attraction."

The purpose of the present pilot study was to investigate the relationship between measures of social attraction and social protection in a sample of emotionally disturbed boys.

## B. METHOD

### 1. Subjects

The subjects were 65 preadolescent boys (ages seven to 13) from The Children's Village, a residential treatment center for emotionally disturbed children. In general, the boys were referred for residential treatment because they were experiencing behavior problems at home (e.g., unmanageable or predelinquent) and/or learning problems at school. Most of the Village boys are from homes wherein, typically, one or more parents are absent, ill, abusive and/or uninterested in the child. The Children's Village does not accept boys who are clearly mentally retarded, have gross neurological or physical handicaps, or who exhibit psychotic behavior. For treatment purposes, the 250 boys in residence at The Village are assigned to one of four comparable units. The subjects in this study represented all the boys residing in one of the four units. The average length of stay of boys at the Village is two years; at the time of the study the median length of stay of the boys tested was 12 months. A total of 34 black, 20 Caucasian, and 11 Puerto Rican boys participated in this study.

### 2. Procedure

*a. Social attraction.* A variant of Kueth's felt figure free placement technique was used to measure social attraction. The Ss were asked to place eight pairs of figures on an  $8\frac{1}{2}'' \times 14''$  sheet of white paper. The figures were paper cutouts with felt backing and ranged in size from 2 to 8 inches. The figures were primarily family members (mother, father, grandmother, grandfather, young boy or self figure, and young sister); however, an older boy friend, a dog, and TV set were also included in the total of nine figures. For the black and Puerto Rican Ss, the figures were all drawn from *The Negro Family* set manufactured by Instructo for flannel boards. The Caucasian Ss were shown comparable figures from *The White Family* set by the same manufacturer. The Ss, tested individually by the author, were read the following instructions:



"Do you see these figures here; if this was your family, who would each one be?" (*S* names them; examiner assists as needed.) "Now I will show you a word that is the name for one of the persons or things in front of you." (*S* is shown name cards for each subject: e.g., mother, grandfather, friend, dog.) "You are to read the word and find the correct person or thing, and then place it on this sheet of paper along with this boy figure who represents yourself. You can place the figures on the paper anywhere you like. I will be tracing the figures so that I can remember later where you placed them." (Tracing the pairs of figures permitted exact measurement of distance between figures to take place at a later time.) The physical distance between self and other figures was taken as an index of the social attraction of the figures to the *S*.

*b. Social protection.* The *Shipwreck Test* was then individually administered to the *Ss* by showing them the shipwreck scene in Figure 1, together with the nine cutout figures described above. The instructions for this test were as follows:

"Let's again pretend that these six figures here are your family; this is you (cutout of a boy about age eight years old), this is your mother, father, grandmother, grandfather, and your sister. This figure is your friend, this is your dog, and this is your TV set. Now look at this picture (Figure 1) of a sinking ship with the six lifeboats alongside. Let's pretend that all of these nine figures are on the ship and that some of them can be saved by getting in the lifeboats. Since only one figure is allowed in each boat, three of the figures will have to stay on the ship. Remember that there is very little chance for anyone who is not in a lifeboat to be saved. Now, who do you want to go in the six lifeboats? Who goes in boat No. 1, 2, 3, etc. . . . ?" The *S*'s six selections from among the nine figures were recorded, and the rank order (1 to 6) in which the figures were assigned to the lifeboats represented the score for that figure on the test. The three figures that were not assigned to a lifeboat were each given the lowest score of 7. Preliminary reliability and validity data for the *Shipwreck Test* are reported in Schaefer (4). The available normative data for the *Shipwreck Test* (5) indicate that normal preadolescent children tend to save all the family figures, while excluding the friend, dog, and television set from the lifeboats.

### C. RESULTS

Table 1 contains descriptive statistics for the measures of social attraction and protection. Inspection of the data for the social attraction scale reveals that the *Ss* placed the grandfather, dog, and friend figures closest to their self figure, while placing the grandmother and television figures fur-

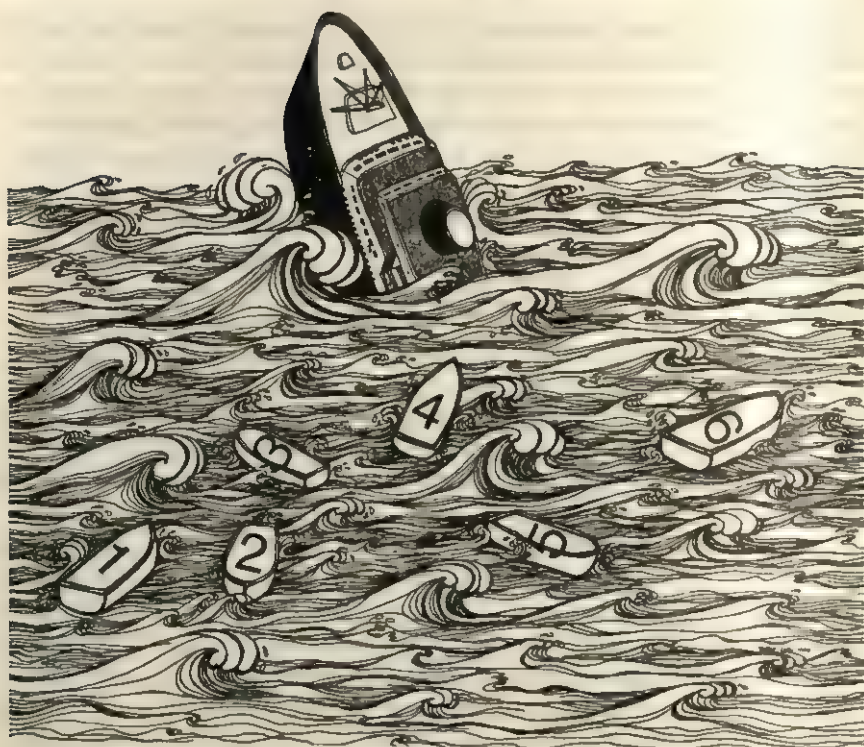


FIGURE 1  
SHIPWRECK SCENE

thetst away. For all nine figures the standard deviations were elevated because of the presence of several extreme scores within each distribution.

Measures of central tendency for the social protection or Shipwreck scale indicate that the three nonfamily members (TV, dog, and friend) were typically excluded from the lifeboats—i.e., rank of 7—while the six family members were usually assigned to one of the lifeboats. Within the family constellation, women and children (mother, grandmother, sister) were usually save first, which is in keeping with usual cultural traditions. The bimodal distributions for the father and grandmother figures indicate that one group of boys assigned these figures a high lifeboat priority, while another group excluded them from the lifeboats.

Further analyses revealed that no significant differences were found between the means of the Caucasian, black, and Puerto Rican groups on either

TABLE 1  
DESCRIPTIVE STATISTICS FOR MEASURES OF SOCIAL ATTRACTION AND SOCIAL  
PROTECTION AND RANK CORRELATION COEFFICIENTS BETWEEN MEASURES

Measure	Mother	Father	Grand- mother	Grand- father	Sister	Friend	Dog	TV	Self
Social attraction <sup>a</sup>									
Mean	4.4	4.5	5.3	3.3	4.3	3.6	3.5	5.4	—
SD	4.6	4.3	4.4	3.4	4.1	3.7	3.4	4.5	—
Social protection									
Median	3.0	4.0	3.0	4.0	3.0	6.0	7.0	7.0	5.0
Mode	1.0	{ 2.0 <sup>b</sup> 7.0 <sup>b</sup>	{ 1.0 <sup>b</sup> 7.0 <sup>b</sup>	4.0	3.0	7.0	7.0	7.0	6.0
Social attraction/ protection <sup>*</sup>									
Spearman rho	.86	.88	.86	.87	.88	.85	.79	.89	

<sup>a</sup> Data for the social attraction scale are in centimeters.

<sup>b</sup> Bimodal.

\* All correlation coefficients significant beyond  $p < .01$ .

the social attraction or social protection scales. Thus, the data for the entire sample were used to investigate the relationship between these two measures.

For this correlation, the physical distance scores on the social attraction scale were converted to relative distances or ranks; i.e., the figure placed closest to the self figure was assigned a rank of 1, the next closest figure was given a rank of 2, etc. The relationship between the ranks assigned to a figure on the two tests—social attraction and social protection—was then determined by Spearman's rank correlation coefficient. These coefficients were found to be significant beyond the  $p < .01$  level for each of the eight figures (see Table 1). Thus, a close relationship was found between these two measures of social regard or orientation.

#### D. DISCUSSION

The major finding of this study is that there seems to be a very close relationship between measures of social attraction and social protection. Thus, the closer a child places a cutout figure of himself to another familiar figure in his home environment, the more protective he tends to be of that figure in an imagined shipwreck situation. These findings lend support to the construct validity of the two measures and indicate that they might be valuable clinical tools for investigating a child's underlying attitudes and feelings toward family members and/or human beings in general. The normative data in this study should contribute toward the identification of children whose scores on the two measures are grossly atypical for this type of population. Further normative studies are needed to investigate the reactions of normal children and to study possible age and sex differences in performance.

An unexpected finding on the social attraction measure was that the Ss tended to show more social attraction for the grandfather, dog, and friend figures than for the parent figures. Since parental neglect, mistreatment, and/or abuse are common in the background history of the Ss, this finding may be a reflection of a deep-seated resentment or hostility by a number of the Ss towards those parental figures who tend to be most closely involved in raising the children: i.e., mother, father, and grandmother. In this regard, Weinstein (7) discovered that emotionally disturbed children tend to construe female parental figures more negatively than do normal children. It would be interesting to contrast the background and personality variables in those children who show less social regard for their parents



on the two measures in this study as opposed to children who exhibit high social orientation toward parental figures on these tests.

Many therapists are becoming more and more inclined to look at their patients against the background of the family unit. The two approaches used in this study for measuring a child's social regard for his family should not only help to give a fuller diagnosis of a child's problem, but even more importantly, they may prove to be of assistance in planning various methods of intervention into family interaction patterns. At the present time there are few useful clinical tools for investigating family dynamics. As Szyrnski (6) noted, it is very difficult to obtain objective data on family dynamics, since there is a natural tendency for the patient to avoid presenting other family members in an unfavorable light. Szyrnski also points out that the well-known mechanism of unconscious repression of unpleasant experiences, particularly if they are threatening an individual's ego ideals, inhibits honest discussion of family interaction patterns. Still other complications in regard to family dynamics arise from the natural difficulty inherent in interviewing young children with limited verbal skills, or hostile, rebellious adolescents. For all these reasons, there is a definite need for simple, easily administered instruments for measuring family dynamics, such as those used in this study. Children tend to perceive both of them as a game and enjoy completing them. It is hoped that this pilot study will encourage future research with these two promising techniques.

#### REFERENCES

1. HEIDER, F. *The Psychology of Interpersonal Relations*. New York: Wiley, 1958.
2. KUETHE, J. L. Social schemas. *J. Abn. & Soc. Psychol.*, 1962, **64**, 31-38.
3. ———. The pervasive influence of social schemata. *J. Abn. & Soc. Psychol.*, 1964, **68**, 248-254.
4. SCHAEFER, C. E. An exploratory study of the relationship between sociometric status and friendship in emotionally disturbed boys. Mimeographed report, The Children's Village, Dobbs Ferry, New York, 1973.
5. ———. Normative data for the *Shipwreck Test*. Mimeographed report, The Children's Village, Dobbs Ferry, New York, 1973.
6. SZYRNSKI, V. Investigation of family dynamics with the "two houses technique." *Psychosomatics*, 1963, **4**, 68-72.
7. WEINSTEIN, L. Social schemata of emotionally disturbed boys. *J. Abn. Psychol.*, 1965, **70**, 457-461.

*The Children's Village*  
Dobbs Ferry, New York 10522





# TIME JUDGMENTS BY MAGNITUDE ESTIMATION AND MAGNITUDE PRODUCTION AND ANXIETY: A PROBLEM OF COMPARISON BETWEEN NORMALS AND CERTAIN SCHIZOPHRENIC PATIENTS\*<sup>1</sup>

*New York State Psychiatric Institute and College of  
Physicians and Surgeons, Columbia University*

---

JACQUES RUTSCHMANN<sup>2</sup>

---

## SUMMARY

The literature suggests association between arousal, general activation, and anxiety on the one hand, and time judgments on the other hand, implying that reported differences in time judgment between nosological groups may be confounded by group differences in arousal-anxiety.

Self-report measures of anxiety, as well as magnitude estimates and magnitude productions of standards ranging from 500 to 2000 msec in 250 msec steps and presented in 10 randomized blocks, were obtained from 16 male normals and from 16 male hospitalized patients with a tentative diagnosis of chronic undifferentiated schizophrenia. Only 10 of the 16 patients were later found to have the same confirmed diagnosis. Data from nine normals and from seven chronic undifferentiated schizophrenics met a criterion of linearity of response functions for both time judgment methods and were further analyzed.

Magnitude estimates and magnitude productions showed underestimation of elapsed time, both types of judgment exhibited satisfactory reliability,

---

\* Recommended by Joseph Zubin of the Editorial Board, received in the Editorial Office on July 24, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was supported in part by Grants M-872 and FR 05560 from the National Institutes of Health, MH 03616 and MH 10315 (USPHS Postdoctoral Research Training-Biological Sciences) from the National Institute of Mental Health, NS 06767 from the National Institute of Neurological Diseases and Stroke, as well as by the N.Y. State Psychiatric Institute where the research was conducted.

<sup>2</sup> The author thanks Dr. Richard D. McConchie for his collaboration in the conceptualization of the study and help with data collection and initial data analysis, and Mr. Maurice Rosengarten for the design, construction, and maintenance of the apparatus. He is indebted to Drs. Lothar Gidro-Frank, Marvin I. Herz, and their staff for help with information and access to patients; to Drs. Joseph L. Fleiss, Maureen B. Kanzler, Richard W. Stewart, and Mr. Phil J. Bowers for help with data analysis and computer programming; and to Drs. L. Erlenmeyer-Kimling and Frank Stallone for their critical comments.

estimates showed "shortening," and productions showed "lengthening" over blocks of trials.

Intercepts of the response  $\times$  standards functions were not generally equal to zero, were more negative in estimation than in production, had marginal reliability or were unreliable, did not correlate significantly between methods, and did not show significant trends over blocks of trials.

Following a model by Carlson and Feinberg, slopes of response  $\times$  standard functions were used as estimates of the rate of the "internal clock." (In estimation, the rate is equal to the slope; in production, it is equal to the reciprocal of the slope.) Average rates of the internal clock did not differ between methods for normals, but were higher in production than in estimation for the seven patients. Clock rates did not differ significantly between groups, were reliable, and exhibited positive correlation between methods. Clock rates exhibited trends over blocks of trials: arctan equivalents of clock rates were linearly related to ordinal numbers of blocks of trials and showed decreases, or "slowing" of the internal clock, in both methods.

Differences in mean anxiety between groups were not significant. In each group, anxiety scores showed positive average correlations with magnitude estimates and negative average correlations with magnitude productions, failed to correlate significantly with intercepts, but showed positive correlations with clock rates.

The data also suggest that anxiety and intrasubject variability may be interrelated.

To conclude: Reported differences in time judgments between nosological groups may not solely be due to nosological differences *per se*, but instead may be due to group differences in anxiety.

---

## A. INTRODUCTION

This report is concerned with comparisons of judgments of short time intervals obtained from normal Ss and from certain schizophrenic patients,<sup>3</sup> and with interrelations between time judgments and anxiety.

It is known that time judgments are influenced by many variables, such as past experience, distraction, monotony, "danger," body temperature, drugs, etc. and that differences in these variables may contribute to differences between time judgments as obtained in different subjects or subject

---

<sup>3</sup> Although it is commonly accepted that the experience of time is altered in various psychopathological states (e. g., 7, 37), time judgments cannot generally be predicted from time experience (e. g., 33, 38).

groups. Goldstone and Goldfarb (22) reported that schizophrenic patients overestimate elapsed time to a greater degree than nonschizophrenic patients or normal Ss and contend that the reported difference may specifically relate to schizophrenia, rather than to chronic illness or to psychiatric disorder in general. Other investigators (e. g., 10, 12) also reported overestimation of elapsed time by schizophrenics.

In the following, an attempt is made to examine further the nosological difference in time judgment and to question the contention that overestimation of elapsed time constitutes a specific disorder for schizophrenic patients.

Time judgments involving semantic relations between words and time intervals (i. e., judgments obtained by the method of verbal estimation or by the method of production) are influenced by verbal and sociocultural norms and show developmental changes (e. g., 18, 23). Time judgments of short standards, as reviewed by Frankenhaeuser (20), are furthermore influenced by changes in partly overlapping and ill-defined variables, such as the rate of flow of mental events, short term memory, attitude, motivation, level of general activity, alertness, fatigue, etc. One possible way to encompass these ill-defined and partly overlapping variables is based on commonly held assumptions regarding arousal (or general activation) in relation to a hypothetical internal clock:<sup>4</sup> time judgments would depend on an internal clock whose rate varies with arousal and activation. In other words, directional changes in one or more measures of arousal or activation (e. g., 28) are assumed to determine directional changes in time judgments.

Two parameters, intercept and slope, specify linear response functions: i. e., functions relating time judgments to the standards that are being judged when a range of standards is used. Here, the rate of the internal clock relative to objective time is determined by the slope of the response function: for verbal estimation, the rate is given by the slope, whereas for production, the rate is given by the reciprocal of the slope.<sup>5</sup> The internal clock may be imagined as a counter counting internal events, with the rate of the internal clock being proportional to the rate of internal events.<sup>6</sup> The internal events

<sup>4</sup> Various authors (e. g., 19, 27, 35) suggested temperature dependent, metabolic, or internal "chemical" clocks. Carlson and Feinberg (10) proposed an event-counting clock, a model to be closely followed in the present report. Other counting clocks were proposed (e. g., 24, 41).

<sup>5</sup> This difference in relations between slope and rate of the internal clock between the two methods arises from the difference between the standards: in verbal estimation, the standards are elapsed times, whereas in production, the standards are verbal labels. It can be seen that if the rate at which verbal estimates grow with standard elapsed times is  $b$ , the rate at which productions grow with standard verbal labels will be  $1/b$ . [For a more detailed presentation of these relationships, see Carlson and Feinberg (10).]

<sup>6</sup> In this model, a particular "tally" in the counter corresponds to a particular "dura-



are probably biochemical or bioelectrical in nature and occur somewhere within the central nervous system. In this way, for example, the overestimation of elapsed time (or "perceptions" that objective time proceeds more slowly, that external time is "dragging") as produced by hyperthyroidism, hyperthermia, administration of dextro-amphetamine, etc. can be understood if one assumes that such effects are mediated by an increase in the rate of internal events.<sup>7</sup>

According to this linear model of the internal clock, equivalence between, for example, "underestimation of elapsed time" and "internal clock slower than external clock" (or "overestimation of elapsed time" and "internal clock faster than external clock") as proposed by Bindra and Waksberg (5) and others, holds only under the assumption that intercepts are zero: i. e., for proportionality between judgments and standards.<sup>8</sup>

While the kind of internal physiological event counted remains as mysterious as the nature of the event-counting clock, the model implies that, other things being equal, only changes in the rate of internal events will result in changes in time judgments. More specifically, if increases or decreases in arousal or activation level are respectively associated with increases and decreases in the rate of internal events, slopes will show direct changes in verbal estimation and inverse changes in production. In other words, higher levels of arousal-activation are expected to lead to a higher rate of internal events and thus to greater slopes in verbal estimation and to smaller slopes in production.

Concerning possible differences in arousal and activation between schizophrenics and normals, the situation is far from clear (e. g., 21, 42, 46);

---

tion word." In verbal estimation, the counter would be reset to zero at the beginning of the standard interval, and the tally reached by the counter at the end of the standard interval would determine the magnitude estimate. In production, the counter would be reset to zero at the onset of a production, and the time taken by the counter to reach the tally specified by the verbal standard would determine the delay between the start and end of the production. The correspondence between tallies and words would have variability, and practice, anchors, knowledge of results, etc. would alter the correspondence between tallies and words.

<sup>7</sup> The linear model of the event-counting clock, as presented, may lack generality: the model may not apply to both "short" and "long" standards, or to such other methods as category scaling, category production, successive comparison, etc. Different or more complicated clocks may be required to account for effects of the sensory modality of the elapsed time and of other concurrent exteroceptive stimulation, and memory variables may play an increasing role in delayed estimation (i. e., judgments of "past time").

<sup>8</sup> The intercept, or changes in intercept, may be interpreted as systematic error of measurement, such as failure to reset the counter at the proper time, failure to reset to zero, constant error in the correspondency between tallies and words, etc. Conceivably, certain independent variables may only affect intercepts, or may affect both intercept and rate.



however, it appears likely that schizophrenics have a higher level of arousal than normals.

Given a frequently voiced general assumption that "anxiety" is associated with arousal and activation,<sup>9</sup> anxiety may be nominated as a possible variable in time judgment tasks: if anxiety acts as a variable of the rate of internal events, it would affect time judgments and thus contribute to variability between nosological groups.

### 1. *Literature Review on Anxiety-Arousal and Time Judgment*

The literature on stress, threat of shock, danger, and on individual differences in self-rating on tests of anxiety suggests the existence of an association between anxiety-arousal and time judgments of standards of less than 60 sec:<sup>10</sup> when relatively strong effects are in evidence, judgments under high anxiety-arousal yield longer verbal estimates (e. g., 25, 34, 39) and shorter productions (e. g., 8, 17, 29) than control conditions. Lack of association (e. g., 9, 13, 40) and occasional reversed effects (e. g., 26) seem related to various conditions which usually increase variability (e. g., unannounced trials in verbal estimation), generate competition between responses (e. g., other activities required from *S* during time judgment tasks), or possibly alter the task in more subtle ways (e. g., instead of magnitude estimation, *S* may perform as in a category scaling experiment).

Stimulus strength, considered to be positively related with arousal, was reported to have the expected effect in verbal estimation (e. g., 4), and spontaneous physiological changes associated with decreases in arousal were found to be associated with longer productions (16).<sup>11</sup>

### 2. *Literature Review on Schizophrenia and Time Judgment*

Adler (1) compared hospitalized schizophrenics, psychopaths, and undergraduate students on several time judgment tasks. Separate analyses of variance indicated that differences between groups were significant for produc-

<sup>9</sup> The relationship between anxiety and arousal is assumed to be asymmetrical: a subject with high anxiety will also show arousal, but changes in arousal and activation may occur in the absence of changes in anxiety.

<sup>10</sup> This restriction is an arbitrary limit and is not meant to imply that processes of time judgment are "uniform" over the range from near zero sec to 60 sec.

<sup>11</sup> The literature on the relationship between EEG and time judgment (e. g., 2, 3, 9, 15, 30, 47, 48) suggests that correlations between time judgments and various estimates of the "number of bioelectric events per unit time" (e. g., apparent alpha rate, percent time alpha, desynchronization, various scores obtained from automatic analysis of the EEG, etc.) are relatively weak, even when time judgment trials and bioelectric recordings were simultaneous.

tion ("S-called" in Adler's terminology), but not for graphic estimation ("E-called"), mean productions being smaller in schizophrenics.

A comparative study by Dobson (14) on normal, neurotic, and time-oriented and time-disoriented schizophrenic World War II veterans resulted in a rather general conclusion that no group differences between mean time judgments were in evidence. Intersubject variability of time judgments was least for the neurotics, with normals and oriented and disoriented schizophrenics showing larger and larger variability.

Four reports (22, 31, 45, 49) are based on a modified method of verbal estimation: Ss reported whether a stimulus was longer or shorter than 1 sec; stimulus durations were changed from trial to trial according to the method of limits. The derived measure, called "second estimation point" (SEP), was the stimulus duration yielding 50 percent shorter and 50 percent longer judgments. These studies indicate that schizophrenics overestimate short intervals of elapsed time when arithmetically spaced step intervals are used in the SEP procedure, as compared to normals, other patients, and schizophrenics in remission. A related study (44) on category scaling indicated that there was no general tendency for schizophrenics to use larger categories than the other groups: schizophrenics used a smaller range of categories than control Ss, and there were indications that this also occurred among more severely ill nonschizophrenic patients. Stated differently, the discrimination of schizophrenics from other subjects only occurred with arithmetically spaced auditory stimuli in the SEP procedure.

In a study comparing verbal estimates and productions (and reproductions) in male normal controls, anxiety neurotics, and chronic undifferentiated schizophrenics (43), the major finding was that nosology is not a variable for mean judgments or for intrasubject variability; however, the interaction (method)  $\times$  (nosology)  $\times$  (duration of standard) reached significance. Intrasubject variability was in general larger for verbal estimation than for production, estimations were significantly larger than productions, and there was a significant negative correlation between verbal estimates and productions ( $Rho = -.51, p < .05$ ).

Carlson and Feinberg (10) studied verbal estimations and productions (and reproductions) in normals and in schizophrenic patients. The duration of a light was used as a standard in estimation (and reproduction), standards ranged from 1 to 10 sec in 1 sec steps, and presentations were random per trial for method and for standard duration.<sup>12</sup> Regression analyses indicated

<sup>12</sup> In other words, the verbal standards in production trials represent the vocabulary for correct verbal estimations, but Ss were not informed of this relationship.

that for 20 out of 20 normal Ss, and for 30 out of 54 schizophrenics, linear trends were significant for each method, without significant deviations from linearity. The average rate of the internal clock for the 30 schizophrenics meeting the criteria for significant linear regression was 1.188 for verbal estimation and 1.322 for production, while for normals it was .842 for verbal estimation and .892 for production. In other words, the internal clock ran faster than objective time in schizophrenics and slower than objective time in normals, but in both groups the clock ran faster in production than in verbal estimation. The relationship between slopes for estimation and production was reciprocal for these schizophrenics, whereas normals showed a linear relationship. Arctan slopes for estimation and production were linearly related within each group: the correlations were  $-.87$  for the 30 schizophrenics and  $-.70$  for normals.

As to intercepts, the situation appears less clear. Theoretically, nonzero intercepts represent systematic errors of measurement. Carlson and Feinberg report the following mean intercepts: Normals, estimation .35 sec, production  $-.36$  sec; 30 schizophrenics, estimation .47 sec, production  $-.01$  sec. However, with the exception of schizophrenics in estimation, coefficients of correlation between intercepts and mean time judgments (mean across standards) were not significant, while correlations between arctan slopes and mean time judgments were sizeable and highly significant. As suggested by Carlson and Feinberg, the meaning of the differences between normals and schizophrenics as reported in their study is unclear because of lack of comparability between the groups.

In summary, among experiments on time judgment in which the performance of schizophrenics was compared to other groups, only Adler (1) permits a rather indirect inference on anxiety—viz., that relative to normals or to psychopaths, the schizophrenics had elevated scores on the "Pt" scale of the MMPI (e. g., 6) and also made shorter productions and (nonsignificantly) longer graphic estimates than normals.

From reports in which no separate assessments of anxiety were made, one gains a mixed impression: Dobson (14) and Warm, Morris, and Kew (43) found no differences between groups; the findings of the Lhamon-Goldstone group vary with method, but suggest that schizophrenics overestimate elapsed time, while Carlson and Feinberg found an accelerated internal clock for schizophrenics as compared to normals.

In the present experiment, a comparison between verbal estimations and productions of chronic undifferentiated schizophrenic patients and of normal volunteer Ss was attempted. For ethical and practical reasons, patients could

not be tested in a drug-free state. Consequently, differences in drug regimen may confound comparisons. However, within the present context, and as a working hypothesis, drugs used in the treatment of schizophrenia may be regarded as variables of anxiety: the essential hypothesis put forward is a relationship between time judgments and anxiety, rather than between time judgments and nosological classification.

### 3. *Hypotheses*

Time judgments and anxiety are related on the assumptions that the rate of hypothetical internal events covaries with anxiety, and that time judgments result from the operation of a hypothetical clock which counts these internal events (see Hypothesis 4.2). As outlined previously, this major hypothesis is based on a linear internal clock and can only be tested by recourse to Ss with linear time judgment functions: Ss whose response *vs.* standard functions can be fitted by straight lines are referred to as "linear Ss," while Ss whose functions cannot be fitted by straight lines are referred to as "nonlinear Ss." With the exception of hypotheses on judgments (see 1.1 through 1.7), all hypotheses were formulated for linear Ss.

Precursory hypotheses on nontransformed time judgments, on parameters (i. e., intercepts and slopes) derived from judgments, as well as on interrelations between judgments and parameters, essentially state expectations based on the literature, on prior experiments, and on congruity between judgments and parameters for normal Ss (see 1.1 through 1.6, 2.1 through 2.4, and 3.1 and 3.2). Other hypotheses concern general expectations for group differences (see 1.7 and 2.5) and empirical questions on trends (see 1.5 and 2.3).

The major hypotheses on interrelations between anxiety and parameters (see 4.2) and the corollary hypotheses on anxiety and judgments (see 4.1) are followed by expectations concerning differences between anxiety measures (see 4.3), group differences in anxiety (see 4.4), and by an empirical question on possible relations between anxiety and trends (see 4.5).

*a. Hypotheses on judgments.* (1.1) Verbal estimations and productions should show reasonable "consistency": i. e., Ss should overestimate (or underestimate) elapsed time with both methods, or else make accurate judgments; there should be substantial positive correlations between judgments and standards for both methods, and judgments should have reliability. (1.2) Intersubject and intrasubject standard deviations of judgments should be larger in production than in estimation. (1.3) Intersubject standard deviations of judgments should correlate with mean judgments and with standards.



(1.4) Mean intersubject standard deviations of judgments, as well as the reliability of judgments, should increase over blocks of trials. (1.5) If there are trends in judgments over blocks of trials,<sup>13</sup> verbal estimates should show "shortening" over blocks of trials, and productions should show "lengthening" over blocks of trials. (1.6) Intrasubject standard deviations of judgments ought to increase with judgments and with standards.<sup>14</sup> (1.7) Intrasubject variability is expected to be larger among schizophrenics than among normals.

*b. Hypotheses on parameters.* (2.1) Intercepts were expected to have zero (or near zero) values and to contain little or no information (e. g., intercepts would not correlate across methods and be relatively unreliable); slopes were expected to have values near  $+1$ , to correlate negatively across methods (i. e., positive correlations between clock rates), and to be reliable. (2.2) Clock rates in production were expected to be higher than clock rates in estimation [see Carlson and Feinberg (10)]. (2.3) On the assumption of the existence of trends of judgments over blocks of trials (see Hypothesis 1.5), the experiment was expected to yield information as to whether trends are reflected as changes in intercept, in slope, or in both parameters. (2.4) While no predictions were made for intercepts, the intrasubject variability of slopes was expected to correlate positively across methods. (2.5) Schizophrenics were expected to have larger intrasubject variabilities than normals for both parameters.

*c. Hypotheses on relations between judgments and parameters.* (3.1) Correlations of mean judgments (across standards) with intercepts ought to be small and variable in sign, while correlations with slopes should be respectable and positive (i. e., correlations of mean judgments with clock rate ought to be positive in estimation and negative in production). (3.2) Intrasubject variability of judgments should correlate positively with intrasubject variability of parameters.

*d. Hypotheses on relations between anxiety and performance measures.* (4.1) Correlations of anxiety measures with mean judgments should be positive for estimation and negative for production. (4.2) Correlations of anxiety

<sup>13</sup> In a previous experiment (32), normal Ss showed lengthening in production with associated increases in relative intersubject variability. However, the same Ss failed to show clear trends of shortening in verbal estimation, and the relative intersubject variability of estimations failed to show clear trends over blocks of trials.

<sup>14</sup> Intrasubject SDs have several theoretical components that cannot be further analyzed by the present experiment: effects of shortening and lengthening are confounded with effects of initial demonstrations of anchors or the dissipation of the effects of demonstrations, with sequential effects within and between blocks of trials, etc. [See also McConchie and Rutschmann (32) and Treisman (41).]



measures with intercepts should be small and variable in sign, whereas correlations with slopes should be positive in estimation and negative in production (i. e., correlations between anxiety and clock rates should be positive for both methods). (4.3) Correlations between anxiety and clock rate were expected to be higher for measures of "state" anxiety than for measures of "trait" anxiety. (4.4) If schizophrenic patients have higher anxiety scores than normals, their clock rate was expected to be higher than the clock rate of normals. (4.5) If there are trends of judgments over blocks of trials (see Hypotheses 1.5 and 2.3), the experiment was expected to determine whether such trends relate to anxiety.

## B. METHOD

### 1. Subjects

Sixteen male inpatients, between the ages of 18 and 36, with a tentative hospital-diagnosis of schizophrenia, chronic undifferentiated type (DSM II, 295.90) were selected, and 16 male medical students served as paid volunteers. However, as explained below, only seven patients and nine normals turned out to meet the requirements for the present study.

Initially, only patients for whom available records failed to reflect a history suggestive of organic brain damage, alcoholism, or drug addiction and patients who had not received somatic therapy within the last three months were asked to volunteer. The number of suitable patients was reduced when a re-examination of hospital charts made approximately 18 months after the experiment revealed that six of the 16 patients either did not in fact meet the initial selection criteria or had a revised diagnosis at the time of testing in the present experiment.<sup>15</sup>

Another reduction in the number of suitable subjects resulted from the findings with tests for the linearity of response functions (see Section C 3a): seven of the 16 normal Ss and three of the 10 chronic undifferentiated schizophrenic patients showed significant departures from linearity and were eliminated from the data analysis. Means and standard deviations of age for the subject groups and subgroups of interest appear in Table 2.

<sup>15</sup> One patient was diagnosed "adjustment reaction of adolescence" (307.2) two days after testing; one was diagnosed "schizophrenia, latent type" (295.5) within six months of testing; two patients received diagnoses of "schizophrenia, paranoid type" (295.3), one within two days of testing and one at an undetermined time prior to the experiment; one was found to have an eight year history of drug abuse (304.0); and one was found to have a history of erythroblastosis neonatorum (744) with multiple transfusions, hypertension, and numerous surgical vascular interventions, with a diagnosis of 295.9 with "mild organic mental brain syndrome."

While none of the elicited reports from normal volunteers contained current use of any drug, most patients were under medication at the time of the experiment. Medication consisted of phenothiazine derivatives, associated with antiparkinsonism drugs or with a tricyclic antidepressant, and in P.R.N. sedatives, as shown in Table 1.

## 2. *Self-Report Measures of Anxiety*

A modified version of the Taylor Manifest Anxiety Scale (Taylor MAS) was used as a measure of "trait" anxiety. The 50 items were administered as questions read by *E*, administered in a fixed sequence. *S* gave "yes" or "no" answers. The "Today Anxiety" scale of the Multiple Affect Adjective Check List (MAACL) was used as a measure of "state" anxiety during the actual time judgment session (50).

## 3. *Apparatus*

The apparatus consisted of a recording section, a time measurement section, a programming section, and a box for *S*.

*a. Recording.* In verbal estimation, the standard interval, and in production, *S*'s production were measured with electronic counters and automatically recorded by a Hewlett-Packard Digital Printer. In verbal estimation, *S*'s judgment was recorded on the printer paper tape by *E*; in production, *E* encoded the verbal standard on the printer tape.

*b. Time measurement.* In verbal estimation, the duration of the standard was measured on each trial, and in production, *S*'s judgment (in the form of the delay between the apparatus initiated onset of a 1000 Hz tone and *S*'s press on a microswitch) was measured (precision  $\pm 1$  msec). Time measurements were made by means of modified Hewlett-Packard AC-4 Digital Counters counting a 1 kHz signal. The shortest measureable time interval was 201 msec.

*c. Programming.* In verbal estimation, each trial began with a presentation of the standard to be verbally estimated: the programming section determined the duration of the standard (precision  $\pm 2$  msec) which was presented to *S* as a 1000 Hz tone at a level of 35 dBSP. The tone was generated by applying a signal from a Hewlett-Packard 206A Audio Signal Generator via an electronic gated audio-amplifier to a loudspeaker. In production, each trial began with a verbally stated standard in msec.

In verbal estimation, *S* was presented with a standard and asked to give his immediate estimate in msec, and *E* recorded each estimate on the printer tape. In production, *S* was asked to produce the verbally stated standard.

TABLE 1  
 MEDICATION AMONG PATIENTS: TOTAL DOSE IN MG PER DAY (VARIABLE DOSE  
 INDICATED IN PARENTHESES SPECIFICS MAXIMUM DAILY DOSE)

Medication	Chronic undifferentiated patients															
	Linear								Nonlinear							
	4	5	8	9	10	12	16	2	6	15	1	3	7	11	13 <sup>a</sup>	14 <sup>b</sup>
Thorazine					600	200	300	150	150	(400)	400	500	(800)	600		b
Mellaril		200		40												
Trilafon					20		40					20			a	
Stelazine	20		3													
Prolixin	8/															
Etrafon	100															
Artane		8	5		8	6	6	4			4	4	(2)	4		b
Cogentin				3												
Seconal																
Chloral											c					
Hydrate <sup>c</sup>																

Note: Each of the 16 patients is designated by a number at the head of column.

<sup>a</sup> Off drugs for one month prior to experiment.

<sup>b</sup> Off drugs for more than two months prior to experiment.

<sup>c</sup> P.R.N. sedative-hypnotic taken the night before the experiment.

Approximately 1 sec after the verbal instruction, the 1000 Hz tone was automatically turned on, and *S*'s task was to switch off the tone (by means of a microswitch) in such a way that the tone had the desired duration. The delay between the closure of the microswitch and its consequences (i. e., turning off the audio gate and stopping the electronic counter) was less than 1 msec.

*d. Box for S.* Attached to the armrest of *S*'s chair was a minibox with a leaf-actuated microswitch.

#### 4. Procedure

The Taylor MAS was administered to patients at the time of the first personal contact, in a quiet office on the ward. Normal volunteers were given the Taylor MAS upon arrival in the laboratory, prior to the short interview and the test session. A short interview, designed to improve communication and to reduce test anxiety was given to all *Ss* before they entered the testing room.<sup>16</sup> After completion of the interview and upon removing time pieces, *S* entered the testing room, a Korfund Audiometric Room supplied with air-conditioning. *E* and the apparatus were located in an adjoining instrument room. An intercom system was used for communications with *S*.

Time judgment methods were administered in a fixed order: after two demonstrations, all trials with the method of verbal estimation were run. Following a rest period, and after four demonstrations, all trials with the method of production were administered. Each *S* judged each of seven standards: i. e., 500, 750, 1000, 1250, 1500, 1750, and 2000 msec with each method. *Ss* were told that the purpose of the examination was ". . . to see how well you can judge time." *Ss* were told that if they gave the same judgment to the same standard, they did better than if they gave different judgments for the same standard, and that they did better if they gave different judgments to different standards (smaller judgments for smaller standards, etc.) no matter how small the difference between the standards.

First, the method of verbal estimation was used. Following explanations (and, if necessary, coaching) and mastery of the millisecond vocabulary to be used, two demonstrations of prelabelled anchors of 1250 msec were given, preceded by the statement that demonstrations help, and that they help more the better one remembers them. Following the demonstrations, *S* was presented with 10 randomized blocks of trials, each block consisting of seven trials, each of the seven standards being presented once in each block. Each

<sup>16</sup> The interview included questions on age, hobbies, sleep, use of cigarettes, coffee, etc., prescription drugs, playing of musical instruments, dancing, preferred TV programs, etc.

trial started with *E* saying, "Ready," followed by the presentation of the standard. The verbal estimation procedure required about 30 minutes.

After a timed five minute rest period spent outside of the laboratory, the experiment progressed to the method of production. Following instructions and another statement as to help to be derived from demonstrations, four demonstrations were given.<sup>17</sup> A prelabelled 1250 msec tone was presented, and *S* was instructed to reproduce this tone. He was told that soon after the 1250 msec tone had gone off, the tone would automatically come on again, and that he was to shut off this second tone by a discrete press on the microswitch in such a way that its duration would be equal to the 1250 msec duration of the first tone. The interval between the demonstration standard and the onset of the reproduction tone was 2000 msec, and the demonstration procedure was run without any feedback or information. Following the four demonstration-reproduction trials, *S* was presented with 10 randomized blocks of production trials, each block consisting of seven trials, each of the seven verbal standards being presented once in each block. As already described, each trial started with a verbal standard (*E* said, "'X' msec. Ready!") followed by the onset of a 1000 Hz tone and, after an interval determined by *S*, by *S*'s response (a single, discrete press on a microswitch) which terminated the tone. This procedure required approximately 20 minutes.

Ss were not told the number of different standards, nor were they told that the same standards were used in both methods. Ss were not told about the randomized block succession of trials and received no feedback or knowledge of results from *E* or from the apparatus.

Upon completing the production task, *S* left the testing room, and after a short rest period, the MAACL was administered in another room. The instructions were "... to describe how you felt while you were inside the testing room, during the time judgment experiment." All Ss were released from the testing session with instructions not to inform potential future Ss of the details of the experiment.

### C. RESULTS AND DISCUSSION

Only data from Ss with linear response functions (see Section C 3 a) for both methods and from those patients who also met the nosological selection criteria (see Section B 1) will be presented: i. e., data from "linear normals"

---

<sup>17</sup> Four trials were considered sufficient to acquaint *S* with the proper operation of the microswitch.



( $N = 9$ ) and from "linear, chronic undifferentiated schizophrenics" ( $N = 7$ ).<sup>18</sup>

Presentation of findings on anxiety, judgments, parameters of linear response functions, relationships between mean judgments and parameters will be followed by the presentation of relationships between anxiety and the various performance measures.<sup>19</sup>

### 1. Anxiety

Anxiety scores were transformed to  $Z$ -scores ( $Z = 10z + 50$ ).<sup>20</sup> For the Taylor MAS, a distribution of college normals ( $N = 233$ ) with a mean of 16 and a standard deviation of 8.1 was used (36). The anxiety scores of the MAACL were transformed by use of a distribution from 44 college students [see Zuckerman and Lubin (50), Table 3, p. 5] obtained with the "Today" form, with a mean of 6.9 and a standard deviation of 3.3. Means and standard

TABLE 2  
AGE DISTRIBUTIONS FOR SUBJECT GROUPS AND SUBGROUPS AND MEANS AND  $SD$ s  
OF  $Z$ -SCORES FOR THE TAYLOR MAS AND THE MAACL TODAY ANXIETY SCALE

Groups	$N$	Age		Taylor MAS		MAACL	
		Mean	$SD$	Mean	$SD$	Mean	$SD$
Linear normals	9	23.8	.8	46.72	10.58	49.67	10.53
All normals	16	23.8	.7	48.94	10.41	53.25	10.58
Chronic undifferentiated patients							
Linear	7	21.2	2.4	54.29	13.87	59.14	10.44
All	10	22.0	—	51.70	12.38	57.00	9.74
All patients	16	22.4	4.2	54.31	13.18	59.56	10.78
All subjects	32	—	—	51.63	12.17	56.41	11.14

Note: Taylor MAS = Taylor Manifest Anxiety Scale; MAACL = Multiple Affect Adjective Check List.

deviations of the resulting distributions of  $Z$ -scores for the groups and subgroups appear in Table 2. The following differences between means of  $Z$ -scores reached significance: between anxiety tests for all patients ( $p < .02$ ), between anxiety tests for combined groups ( $N = 32$ ,  $p < .005$ ). Specifically, differences between patients and normals were not significant, nor were differences between the respective linear subgroups.

<sup>18</sup> Other subgroups were "nonlinear normals" ( $N = 7$ ); "nonlinear, chronic undifferentiated schizophrenics" ( $N = 3$ ); "linear other psychiatric patients" ( $N = 5$ ); "nonlinear other psychiatric patients" ( $N = 1$ ). No attempt was made to further analyze nonlinear functions. Carlson and Feinberg (10) discussed possible analyses of nonlinear functions, but concluded that larger numbers of nonlinear  $S$ s would be needed for that purpose.

<sup>19</sup> All correlation coefficients in this section are Pearson  $r$ s.

<sup>20</sup> Here  $z$  has the usual meaning of  $(X - \bar{X})/SD$ .

Correlations between the tests were  $+ .61$  ( $p < .05$ ) for normals,  $+ .84$  ( $p < .01$ ) for patients, and  $+ .75$  ( $p < .01$ ) for the combined groups. For the linear subgroups, the coefficients were  $+ .56$  (NS) for the nine normals and  $+ .87$  ( $p < .05$ ) for the seven schizophrenics.

The interpretation of differences in mean anxiety measured by the two tests is no doubt complex: given that unrelated distributions were used to effect the transformation to  $Z$ -scores, one cannot infer that, for example, differences in contributions of "test anxiety" generated the observed differences, with patients showing "significant" test anxiety in the MAACL test, etc. Correlations between the tests essentially confirm previous reports (e. g., 50) and fall near the highest possible coefficients predictable on the basis of published reliability coefficients of the tests.

The absence of significant differences between group means in anxiety was unexpected and prevents testing Hypothesis 4.4. Possible *ex post facto* explanations for the lack of group differences may involve drug effects (see Table 1) that may have reduced anxiety among schizophrenics<sup>21</sup> and stress on normal Ss (medical students) resulting from their condition of life.

## 2. Time Judgments

Judgments were submitted to separate ANOVAS for each linear subgroup and each method, with use of a mixed model with "Standards" and "Blocks" as fixed variables. Table 3 gives a summary of these analyses from which it can be seen that standards and blocks are significant sources of variance. Group means of individual mean judgments (for each method and each standard, a mean for each  $S$  was obtained by averaging the 10 judgments, one from each block) appear to increase approximately linearly with standards as shown in Table 4. This main effect is reflected in average correlations between standards and mean judgments (correlations were computed between standards and mean judgments across subjects for each block and averaged across blocks via Fisher's  $r$  to  $Z$  transformation): Normals, estimation  $+ .99$ , production  $+ .99$ ; Patients, estimation  $+ .99$ , production  $+ .98$ , each average coefficient being significant beyond the .001 level. Table 4 also shows that in estimation, mean judgments are smaller than the standards and that in production, mean judgments are larger than the standards; i. e., both methods reveal underestimation of elapsed time, a consistency expected in Hypothesis 1.1.

---

<sup>21</sup> Correlations between anxiety  $Z$ -scores and dose levels expressed relative to chlorpromazine were not significant: for the Taylor MAS, the coefficient was  $-.24$ ; for the MAACL Today Anxiety scale, it was  $-.21$ .

TABLE 3  
ANOVAS FOR EACH GROUP AND EACH METHOD ON JUDGMENTS (ONLY LINEAR  
Ss ARE INCLUDED)

Variables	Normals				Patients			
	df	SS	MS	F	df	SS	MS	F
<i>Magnitude estimates</i>								
Subjects (Ss)	8	57,164,331	7,145,541		6	28,742,712	4,790,452	
Standards (Ts)	6	135,311,425	22,551,904	80.81***	6	72,395,622	12,065,937	83.01***
Blocks	9	7,997,625	888,625	6.18***	9	4,859,546	539,950	7.02***
Ss X Ts	48	13,394,999	279,062		36	5,232,806	145,356	
Ss X Blocks	72	10,360,664	143,898		54	4,153,615	76,919	
Blocks X Ts	54	3,395,022	62,871	1.48*	54	3,321,240	61,504	1.58**
Blocks X Ts X Blocks	432	18,339,362	42,452		324	12,626,224	38,970	
Total	629	245,963,427			489	131,331,765		
<i>Magnitude productions</i>								
Subjects (Ss)	8	215,041,633	26,880,204		6	104,274,162	17,379,027	
Standard (Ts)	6	202,624,161	33,770,694	30.12***	6	117,165,231	19,527,538	32.47***
Blocks	9	10,294,915	1,143,879	3.72**	9	8,351,914	927,990	3.54**
Ss X Ts	48	53,818,103	1,121,210		36	21,649,228	601,367	
Ss X Blocks	72	22,128,151	307,335		54	14,161,256	262,245	
Blocks X Ts	54	5,372,162	99,484	1.48*	54	3,956,149	73,262	1.02(NS)
Blocks X Ts X Blocks	432	29,123,485	67,415		324	23,372,957	72,139	
Total	629	538,402,611			489	292,930,896		

\*  $p < .05$ .  
 \*\*  $p < .01$ .  
 \*\*\*  $p < .001$ .

TABLE 4  
GROUP MEANS (ACROSS LINEAR Ss) OF INDIVIDUAL MEAN JUDGMENTS (ACROSS BLOCKS) AND OF INTRASUBJECT SDs AND MEAN (ACROSS BLOCKS) INTERSUBJECT SDs FOR EACH STANDARD, AS WELL AS MEANS (ACROSS STANDARDS) OF GROUP MEANS OF JUDGMENTS AND RESPECTIVE MEANS (ACROSS STANDARDS) OF INTERSUBJECT SDs FOR EACH BLOCK (IN MSEC)

Fixed Variable	Estimation						Production					
	Normals			Patients			Normals			Patients		
	Judgment	Intra-S	Inter-S	Judgment	Intra-S	Inter-S	Judgment	Intra-S	Inter-S	Judgment	Intra-S	Inter-S
Standard												
500	289	80	160	254	113	163	672	147	259	788	217	353
750	523	169	233	459	217	244	937	198	359	1050	249	363
1000	777	183	296	679	217	321	1168	195	499	1245	253	436
1250	988	263	345	851	232	420	1471	217	634	1458	300	489
1500	1232	273	463	1037	268	356	1766	312	845	1653	330	620
1750	1431	330	541	1234	273	395	2027	355	993	1944	341	740
2000	1694	323	601	1411	229	376	2389	403	1092	2323	363	943
Block												
1	1252		322	1082		296	1259		399	1207		342
2	1140		483	973		315	1319		485	1394		659
3	986		306	857		318	1429		523	1379		562
4	1002		329	859		260	1390		555	1501		520
5	868		330	809		345	1497		630	1560		500
6	911		455	746		345	1559		740	1536		470
7	896		306	765		348	1598		821	1513		525
8	925		336	797		298	1583		815	1641		733
9	976		400	780		341	1653		888	1682		773
10	950		501	795		381	1612		831	1532		550

Means (across standards) of group mean judgments for each block are also listed in Table 4. It can be seen that the main effect of blocks is an approximately linear decrease of mean estimates over blocks (i. e., shortening) and an approximately linear increase in mean productions over blocks (i. e., lengthening). Average correlations between the ordinal number of blocks and mean judgments (correlations were computed between ordinal numbers and respective mean judgments across Ss for each standard, and averaged across standards via Fisher's  $r$  to  $Z$  transformation) were as follows: Estimation, normals  $-.60$ , patients  $-.66$ ; Production, normals  $+.84$ , patients  $+.70$ , each average coefficient being significant beyond the .001 level. These trends over blocks of trials are consistent with Hypothesis 1.5.

As found by the separate ANOVAS, the standard  $\times$  block interactions (which account for at most 2.5 percent of the total variance) reached significance in three out of four analyses. Graphic plots of interactions (not shown) suggest more pronounced trends for longer standards than for shorter standards.

Table 4 also lists mean (across Ss) intrasubject standard deviations, as well as mean (across blocks) intersubject standard deviations for each standard, and mean (across standard) intersubject standard deviations for each block.

The existence of trends of judgments over blocks of trials restricts the meaning and representativeness of mean judgments across blocks, of intersubject, as well as of intrasubject variability. On the assumption that there are linear trends, the greater the number of blocks of trials in an experiment, the smaller the mean estimation, the larger the mean production, and the larger the intrasubject standard deviations. Also, estimates of reliability may be affected by trends and may be higher than "true" reliability. In spite of the fact that the data of the present experiment may not meet all prerequisites, statistics on variability and reliability, as well as intercorrelations, will be presented below. In a later section, the recourse to parameters of linear response functions will permit a more rigorous analysis (see Section C 3).

Grand means of intersubject standard deviations (across standards and across blocks) were as follows: Normals, estimation 377 msec, production 563 msec; Patients, estimation 325 msec, production 563 msec. Grand means of intersubject coefficients of variation were as follows: Normals, estimation 40.9%, production 43.9%; Patients, estimation 44.4%, production 37.3%. Within each group, differences between the standard deviations for the two methods were significant ( $p < .001$ ), and as predicted by hypothesis 1.2,



productions exhibited significantly greater standard deviations than estimations. Differences between coefficients of variation were not significant. Differences between groups failed to reach significance for both standard deviations and coefficients of variation. As predicted by Hypothesis 1.3, correlations between standards and mean intersubject standard deviations (averaged across blocks at each standard) were positive: Normals, estimation  $+ .99$ , production  $+ .99$ ; Patients, estimation  $+ .82$ , production  $+ .96$ , all coefficients being significant beyond the  $.01$  level. The corresponding correlations for mean coefficients of variation were as follows: Normals, estimation  $- .79$  ( $p < .01$ ), production  $+ .87$  ( $p < .01$ ); Patients, estimation  $- .97$  ( $p < .01$ ), production  $- .15$  (NS).

Correlations between mean judgments and corresponding intersubject variability were computed for each block, and averaged over blocks via Fisher's  $r$  to  $Z$  transformation. As expected (Hypothesis 1.3), the average correlations between mean judgments and standard deviations were relatively high: Normals, estimation  $+ .92$ , production  $+ .96$ ; Patients, estimation  $+ .69$ , production  $+ .90$ , all average coefficients being significant beyond the  $.001$  level. Corresponding average correlations involving coefficients of variation (for which no predictions were made) were as follows: Normals, estimation  $- .56$  ( $p < .001$ ), production  $+ .54$  ( $p < .001$ ); Patients, estimation  $- .71$  ( $p < .001$ ), production  $+ .03$  (NS).

Correlations between mean intersubject variability and ordinal number of blocks (mean variability being obtained by averaging standard deviations or coefficients of variation over standards from the same block) were as follows for standard deviations: Normals, estimation  $+ .30$  (NS), production  $+ .97$  ( $p < .01$ ); Patients, estimation  $+ .59$  (NS or  $p < .10$ ), production  $+ .50$  (NS). Correlations for mean coefficients of variation were as follows: Normals, estimation  $+ .69$  ( $p < .05$ ), production  $+ .93$  ( $p < .01$ ); Patients, estimation  $+ .65$  ( $p < .05$ ), production  $+ .20$  (NS). While several coefficients failed to reach statistical significance, all turned out to be positive, and the prediction of an increase in mean intersubject standard deviations over blocks of trials for normals (Hypothesis 1.4) is, as in a previous experiment (32), valid for production, but not for estimation.<sup>22</sup> However, contrary to the earlier finding, coefficients of variation in normals significantly increased over blocks of trials in both methods.

The reliability of judgments was estimated by means of correlation coefficients between successive pairs of blocks (i. e., block 1 *vs.* block 2, block 3 *vs.*

<sup>22</sup> For estimation, and given shortening, one would predict correlations with coefficients of variation to be larger than correlations involving standard deviations, while for productions, given lengthening, one would expect the inverse.

block 4, etc.) at each standard. Mean reliabilities for each pair of blocks were obtained by averaging across standards ( $N = 7$ ) via Fisher's  $r$  to  $Z$  transformation, and an overall estimate was obtained by averaging mean reliabilities over successive pairs of blocks ( $N = 5$ ). Overall reliabilities were as follows: Normals, estimation .77, production .80; Patients, estimation .91, production .83. The reliabilities for block 1 *vs.* 2 were slightly lower than for later pairs of blocks; however, contrary to hypothesis 1.4, there was no regular progressive increase in reliability after block 2 or 3.

Mean (across  $Ss$ ) intrasubject standard deviations did not differ between methods for either group, and differences between groups were not significant for either method—findings that are at variance with Hypothesis 1.7 and that are not consonant with Hypothesis 1.2 for intrasubject standard deviations. Mean intrasubject coefficients of variation exhibited differences between methods: for each group, the coefficients were significantly larger in estimation than in production, but differences between groups failed to reach significance (see Hypothesis 1.7).

Correlations between mean judgments and mean intrasubject standard deviations were as follows: Normals, estimation  $+ .75$  ( $p < .05$ ), production  $+ .90$  ( $p < .01$ ); Patients, estimation  $+ .38$  (NS), production  $+ .85$  ( $p < .05$ ), a finding that is consonant with Hypothesis 1.6. For normals, correlations between standards and mean intrasubject standard deviations were  $+ .96$  ( $p < .01$ ) in estimation and  $+ .97$  ( $p < .01$ ) in production, while in patients, the corresponding coefficients were  $+ .74$  (NS) and  $+ .99$  ( $p < .01$ ). Correlations between intrasubject standard deviations in estimation *vs.* production were computed between means (across standards) as well as separately for each standard: the correlation coefficients were variable in magnitude and in sign, but there were indications that the association might be negative.

Correlations of mean relative intrasubject variability with mean judgments were not significant, while correlations with standards were as follows: Normals, estimation  $- .85$  ( $p < .01$ ), production  $- .69$  (NS); Patients, estimation  $- .96$  ( $p < .01$ ), production  $- .97$  ( $p < .01$ ). Correlations between relative intrasubject variability in estimation *vs.* production were variable in magnitude and in sign, with indications that the association might be positive.

### 3. Linear Response Functions

a. *Parameters.* Straight lines (i. e., regression lines of the form  $Y = A + B_{(YX)}X$ , where  $Y$  stands for judgment and  $X$  for standard) were fitted to the data of each  $S$ , separately for each method. For each  $S$  and each method, 10

individual regression lines (one for each block), nine regression lines for pairs of blocks (i. e., blocks 1-2, blocks 2-3, . . . blocks 8-9, blocks 9-10), and one overall regression line (for blocks 1 through 10) were obtained by the method of least squares. For each *S* and method, a test for the existence of overall (block 1 through 10) linearity was made by comparing the variance of means about the overall regression line ( $df = 5$ ) with the "adjusted" variance about the means, the latter being obtained by subtracting from the within standards sum of squares ( $df = 63$ ) the sum of squares of block means about the overall mean and of block slopes about the overall slope [reduction in  $df = 2(10-1)$ ], thus yielding an *F* ratio with 5 and 45 degrees of freedom. Similar *F* tests (with 5 and 5 degrees of freedom) were made for pairs of blocks.

With the goal of tolerating some overall departure from linearity in the absence of significant departures from linearity for pairs of blocks [so as to afford some degree of comparability with the Carson and Feinberg (10) data], a composite criterion was adopted to classify *Ss*: a subject was classified as "nonlinear" when the level of significance of the overall (block 1 through 10) *F* value for departure from linearity reached the .001 level in one of the methods, or, in case the overall *F* value reached the .01 level but not the .001 level, when one or more of the *F* values computed for pairs of blocks for the same method reached the .05 level of significance. *Ss* not classified as nonlinear by this procedure were regarded as "linear." This resulted in nine linear normals and seven linear, chronic undifferentiated schizophrenics (see section B 1).

Means across *Ss* of mean intercepts and mean slope coefficients (mean over blocks 1 through 10) are given in Table 5, which also lists intersubject standard deviations for the linear subgroups of interest. Mean intercepts differed significantly from zero for normals in estimation and for patients in production (in each case  $p < .001$ ), but not for normals in production and for patients in estimation. Differences between mean intercepts in estimation and production were significant for both linear groups. Correlations between intercepts in estimation and production were not significant (for normals, the coefficient is  $-.54$ , for patients  $+.01$ ). Group differences in mean intercept were not significant for estimation, while for production, the difference of  $-205$  msec was significant beyond the .05 level.

For each subject and each method, the mean intercept was based on 10 intercepts, one from each block. Correlations between such means and the respective standard deviations (i. e., "absolute" intrasubject variability) were not significant.

TABLE 5  
 MEANS (ACROSS Ss) OF OVERALL INTERCEPTS, OF OVERALL SLOPES, OF MEAN ARCTAN VALUES REPRESENTATIVE OF  
 CLOCK RATES, AND OF CLOCK RATES CORRESPONDING TO MEAN ARCTAN VALUES

Group or subgroup	N	Estimation			Production				
		Intercept (msec)	Slope (tangents)	Arctan slope (degrees)	Clock rate (tangents)	Intercept (msec)	Slope (tangents)	Arctan value (degrees)	Clock rate (tangents)
Linear normals									
M	9	-167	.9264	41.32	.8792	+ 75	1.1327	45.50	1.0176
SDa		91		8.40		75		13.57	
Nonlinear normals									
M	7	+ 31	.8154			- 17	1.1276		
All normals									
M	16	- 80	.8779			+ 34	1.1304		
Chronic undiff. linear patients									
M	7	-113	.7683	36.70	.7454	+280	.9716	48.36	1.1251
SDa		116		5.74		77		9.86	
All chronic undiff. patients									
M	10	- 97	.7964			+214	1.1052		
Other patients									
M	6	-197	1.0723			+211	1.0000		
All patients									
M	16	-134	.8999			+213	1.0095		

Note: In estimation the value is the arctan slope; in production the value is 90° minus the arctan slope.

a. Intersubject SDs.



The reliability of intercepts was assessed by means of correlations between intercepts from successive pairs of blocks (i. e., block 1 *vs.* 2, block 3 *vs.* 4, . . . block 9 *vs.* 10). Mean reliabilities over successive pairs of blocks (i. e., over five correlations) were obtained via Fisher's  $r$  to  $Z$  transformation. Mean reliabilities for intercepts were as follows: Normals, estimation  $+ .21$  (NS), production  $+ .54$  ( $p < .01$ ); Patients, estimation  $+ .51$  ( $p < .05$ ), production  $+ .36$  (NS). Thus, the findings on intercepts are not as simple as predicted by Hypothesis 2.1.

Analyses of slopes were based on arctan equivalents (e. g., the slope of  $+1$  has an arctan equivalent of  $+45.0^\circ$ ), since distributions of slope coefficients may be presumed to be skewed. For arctan slopes, the mean reliabilities (derived over successive blocks of trials, etc.) were as follows: Normals, estimation  $+ .81$ , production  $+ .92$ ; Patients, estimation  $+ .61$ , production  $+ .81$ , each mean reliability coefficient being significant beyond the .01 level. Here, it may be noted that while reliability is present in judgments (see Section C 2), among parameters of linear regression, and as predicted by Hypothesis 2.1, only the slopes (expressed as arctan slopes) exhibit adequate reliability, whereas intercepts are unreliable or show only marginal reliability.

Mean (across Ss) overall arctan slopes (obtained separately for each method by transforming a subject's slope coefficients for individual blocks to arctan equivalents, obtaining a mean arctan slope for each S, and averaging such means across Ss) appear in Table 5. For estimation, the mean arctan slope and the equivalent tangent are listed, while for production, the complement of the mean arctan slope (i. e.,  $90^\circ$  minus mean arctan slope) and its equivalent tangent are listed. For estimation, the equivalent tangent is an estimate of the ratio of the rate of the internal clock relative to objective time, whereas in production, the tangent equivalent of the complement of the mean arctan slope is a measure of the same ratio. Overall arctan slopes were significantly different from zero in both methods for both linear subgroups of interest ( $p < .001$ ).

Differences between clock rates were tested by means of  $t$  tests on arctan transformed slopes for estimation, and on complements of arctan slopes for production. Differences between methods were significant for patients ( $p < .01$ ), but did not reach significance for normals. Thus, in linear, chronic undifferentiated schizophrenics, the internal clock ran significantly faster in production than in estimation (e. g.,  $1.125$  *vs.*  $.768$ ), but contrary to Hypothesis 2.2, in normals, the difference between  $1.018$  and  $.879$  was not significant. In both groups, clock rates between methods showed significant positive correlation as predicted by Hypothesis 2.1: in normals, the correlation was  $+ .68$ ; in patients it was  $+ .83$  (each coefficient exceeds the .05 level of



significance). Differences between mean clock rates of normals and patients were not significant.

For each *S* and each method, the mean arctan slope was a mean of 10 arctan slopes, one from each block. Correlations between such means and the respective standard deviations (i. e., "absolute" intrasubject variability of arctan slopes) were not significant.

Parameters from individual blocks of trials were further examined, since means across *Ss* of overall (mean over blocks 1 through 10) parameters and their variability may not adequately describe the findings. Intercepts for individual blocks showed variation: in estimation, the extreme intercepts were -645 and 170 msec among normals and -690 and 386 msec among patients, whereas in production, the extreme values were -642 and 686 msec for normals, and -300 and 1140 for patients. Standard deviations (i. e., absolute intrasubject variability) of intercepts in estimation ranged from 101 to 225 msec (mean 139.4 msec) among normals and from 71 to 282 msec (mean 179.7 msec) among patients, the corresponding values in production being 106 to 521 msec (mean 197.2 msec) among normals and 114 to 450 msec (mean 231.3 msec) among patients. Differences between means of intrasubject standard deviations of intercepts—i. e., between methods within groups, or between groups within methods—were not significant. With use of a level of significance of .05, seven out of nine normals and four out of seven patients had nonzero overall intercepts in estimation, while six out of nine normals and six out of seven patients had nonzero overall intercepts in production. (In estimation, the nonzero intercepts were negative, while in production, four of the six nonzero intercepts among normals were positive, and all nonzero intercepts among patients were positive.) Correlations between intrasubject standard deviations and overall intercepts were not significant. Intercorrelations between intrasubject standard deviations of intercepts between estimation and production were not significant.

Slope coefficients for individual blocks were tested for differing significantly from a slope of zero by testing the significance of correlation coefficients between judgments and standards, separately for each subject, each method, and each block. For linear normals, 89 out of 90 coefficients in estimation and 87 out of 90 in production were significantly different from zero ( $p < .05$ ). For linear patients, 69 out of 70 coefficients in estimation and 64 out of 70 coefficients in production were significantly different from zero ( $p < .05$ ). All slope coefficients for individual blocks were positive for both methods and for both groups. For each subject, overall (block 1 through 10) slopes were significantly different from zero (in each case,  $p < .001$ ).

Standard deviations ("absolute" intrasubject variability) of individual

overall arctan slopes in estimation ranged from  $3.5^\circ$  to  $8.2^\circ$  (mean  $5.32^\circ$ ) in normals and from  $2.3^\circ$  to  $7.8^\circ$  (mean  $5.15^\circ$ ) in patients, the corresponding values for production being  $3.1^\circ$  and  $7.5^\circ$  (mean  $5.05^\circ$ ) among normals and  $4.2^\circ$  and  $8.8^\circ$  (mean  $6.36^\circ$ ) among patients. Coefficients of variation (i. e., "relative" intrasubject variability expressed as a percentage of the overall arctan value corresponding to the clock rate) in estimation ranged from 7.8% to 33.0% (mean 14.0%) in normals and from 6.1% to 22.2% (mean 14.2%) in patients, while for production, the values were 6.0% to 30.6% (mean 12.3%) in normals and 8.8% to 22.2% (mean 13.8%) in patients. None of the differences between mean absolute or relative variability of arctan values—i. e., between methods within groups, or between groups within methods—were significant. Thus, neither parameter showed the larger intrasubject variability for schizophrenics as predicted by Hypothesis 2.5.

Correlations between intrasubject standard deviations of arctans and overall arctan values were not significant. Contrary to Hypothesis 2.4, correlations between intrasubject standard deviations of arctan slopes in estimation and production were not significant: normals .51 (NS), patients .59 (NS). For corresponding coefficients of variation in terms of arctan values representative of clock rates, the correlations were as follows: normals .89 ( $p < .01$ ), patients .39 (NS).

Correlations between intrasubject standard deviations of intercepts and of arctan slopes were as follows: Normals, estimation .01 (NS), production .65 (NS); Patients, estimation .72 (NS), production .86 ( $p < .05$ ).

*b. Trends of parameters over blocks.* Statements in the literature on shortening in estimation and lengthening in production imply that one or the other or both parameters change systematically with blocks. Given Falk and Bindra's (17) findings on apparently linear lengthening of productions of 15 seconds with ordinal number of blocks of trials,  $F$  tests for the existence of linear regression were made separately for each subject group and each method for functions relating arctan slope to ordinal number of blocks, as well as for intercepts *vs.* ordinal number of blocks. The variance of mean values (intercepts or arctan slopes) about the regression line ( $df = 8$ ) was compared with the "adjusted" variance about block means, the latter being obtained by subtracting from the within blocks sum of squares [ $df = 10 (N - 1)$ , with  $N$  equal to the number of subjects] the sum of squares of subject means about the overall mean and the subject rate of change per block about the overall rate of change per block [reduction in  $df = 2 (N - 1)$ ]. Neither intercepts nor slopes exhibited significant departures from linearity.

The rate of change of intercepts *vs.* blocks or of arctan slopes *vs.* blocks

was tested for being significantly different from a rate of change of zero by means of the  $t$  test for significance of the correlation coefficient between the parameter (i. e., intercept or slope) and blocks. For intercepts and for estimation, two normals and one patient exhibited rates of change that were significantly different from zero ( $p < .05$ ): in one normal, the intercept increased at a rate of 32 msec per block, in the other at 52 msec per block, while in the patient, the intercept decreased at a rate of 36 msec per block. In production, one normal and one patient showed significant ( $p < .05$ ) nonzero rates: the normal had a decrease in intercept of 33 msec per block, while the patient showed an increase of 44 msec per block. However, mean rates of change of intercepts across Ss (see Table 6) did not significantly differ from zero, did not significantly differ between estimation and production for either group, and did not significantly differ between normals and patients for either method (in each case,  $p > .10$ ). Correlations between overall intercepts and rate of change of intercepts per block were not significant.

With respect to rates of change of arctan slopes *vs.* blocks, three normals and two patients showed significant negative rates of change in estimation, while in production, three normals and three patients showed significant positive rates of change ( $p < .05$ , one-tailed test). However, most of the individually nonsignificant rates of change agreed in their sign (i. e., decrease in slope over blocks in estimation, increase in slope over blocks in production) such that the mean rates of change of slopes across Ss were significantly different from zero (in each case,  $p < .05$  or better by one-tailed tests, see Table 6). Thus, the present experiment indicates that shortening in estimation and lengthening in production over blocks of trials are linked to variations in slope, rather than to variations in intercepts or in both parameters (see Hypothesis 2.3). More specifically, linear decreases (or increases) in arctan slopes over blocks of trials describe the shortening in estimation (or the lengthening in production).

Differences between mean rates of change of arctan slopes between estimation and production were significant (for normals  $p < .001$ , for patients  $p < .01$ ). Differences between mean rates of change of arctan values representative of clock rates, however, were not significant, suggesting that the "slowing" of the internal clock is the same in both methods.<sup>23</sup> Correlations between the rate of change of arctan values per block in estimation *vs.* production

---

<sup>23</sup> This finding is consistent with a recent report by Carlson and Feinberg (11): within session variations in estimation and production exhibited changes toward slower counting rates, the changes being consistent across methods, but the report does not take issue with the nature of within session trends of clock rates.

TABLE 6  
FUNCTIONS FOR LINEAR PARAMETERS *vs.* BLOCKS: MEAN ORDINATES FOR BLOCK ZERO AND MEAN RATE  
OF CHANGE PER BLOCK FOR INTERCEPTS AND FOR ARCTAN SLOPES

Groups	Intercepts <i>vs.</i> Blocks			Arctan Slopes <i>vs.</i> Blocks		
	Estimation	Production		Estimation	Production	
	Ordinate block zero (msec)	Ordinate block zero (msec)	Rate per block (msec/ block)	Ordinate block zero (degrees)	Ordinate block zero (degrees)	Rate per block (deg/ block)
Linear normals ( $N = 9$ )	-208	+7.3	-15.2	45.8	40.3	+77***
Linear chronic undiff. patients ( $N = 7$ )	-88	-4.8	+5.2	40.0	37.9	+68**

<sup>a</sup> One-tailed. All other probabilities based on two-tailed tests.

\*  $p < .05$ .

\*\*  $p < .01$ .

\*\*\*  $p < .001$ .



were not significant. Differences between normals and patients in mean rate of change of arctan values were not significant for either method. Correlation coefficients between overall arctan values and rate of change of arctan values per block were as follows: Normals, estimation  $+ .22$  (NS), production  $+ .69$  ( $p < .05$ ); Patients, estimation  $+ .04$  (NS), production  $+ .26$  (NS). Correlations between intrasubject standard deviations of arctan slopes and the rate of change of arctan values per block were not significant.

With regard to the description of findings on Ss with linear response functions, means across Ss of overall intercepts as listed in Table 5 are representative in the sense that there is no significant trend of intercepts as a function of blocks of trials. However, overall arctan values representative of clock rates, or their tangents, only furnish an interpolated rate that applies between blocks 5 and 6. In consideration of arctan slopes, and given the mean rate of arctan slopes per block as listed in Table 6, it can be seen that with an increase in the ordinal number of the block, arctan slopes of response functions decrease in estimation (shortening) and increase in production (lengthening), both average changes being linear over the 10 blocks of trials of the present experiment. In terms of clock rates associated with arctan slopes, both methods reveal "slowing" of the internal clock with ordinal number of blocks (see Hypothesis 2.3).

#### 4. *Relationship Between Mean Judgments and Parameters*

Correlation coefficients between grand means of judgments (across blocks and across standards) and overall (block 1 through block 10) intercepts were as follows: Normals, estimation  $-.24$  (NS), production  $-.48$  (NS); Patients, estimation  $.59$  (NS), production  $-.33$  (NS). Correlations between grand mean judgments and overall arctan values representative of clock rates were as follows: Normals, estimation  $.93$ , production  $-.89$ ; Patients, estimation  $.87$ , production  $-.96$  (each coefficient being significant at, or beyond, the .01 level). These findings are consonant with Hypothesis 3.1.

Correlations between intrasubject variability of judgments and the corresponding intrasubject variability of parameters were as follows: Between standard deviations, judgments *vs.* intercepts, Normals, estimation  $.71$  ( $p < .05$ ), production  $.98$  ( $p < .01$ ); Patients, estimation  $.75$  ( $p < .10$ ), production  $.32$  (NS). Between standard deviations, judgments *vs.* arctan slopes, Normals, estimation  $.41$  (NS), production  $.53$  (NS); Patients, estimation  $.72$  ( $p < .10$ ), production  $.31$  (NS). Between coefficients of variation of judgments and relative variability of arctan slopes in terms of arctan values representative of clock rates, Normals, estimation  $.90$  ( $p < .01$ ), production



.59 ( $p < .10$ ); Patients, estimation .80 ( $p < .05$ ), production .47 (NS). For normals, the association between standard deviations of judgments and standard deviations of intercepts, as well as the correlations between coefficients of variation of judgments and relative variability of arctan slopes, are consonant with Hypothesis 3.2, while correlations between standard deviations of judgments and standard deviations of arctan slopes failed to reach significance.

### 5. *Anxiety and Performance Measures*

As outlined in Section C 1, the present experiment did not yield group differences in anxiety as measured with the Taylor MAS or the Today Anxiety scale of the MAACL. Consequently, the following paragraphs are addressed to questions on relationships between anxiety and time judgments within each group.

*a. Anxiety and judgments.* Correlation coefficients were computed between each anxiety  $Z$ -score and mean judgments (mean across blocks) for each linear group, each method, and each standard. Average correlations across standards, obtained via Fisher's  $r$  to  $Z$  transformation, are listed in Table 7. It can be seen that significance is present for normals for the Taylor MAS for both methods, which conforms to the predictions of Hypothesis 4.1. For patients, significance is limited to the MAACL in estimation. Average correlations between anxiety  $Z$ -scores and intrasubject variability of judgments (i. e., standard deviations or coefficients of variation) also appear in Table 7. The table suggests that among normals, the higher the Taylor MAS score, the larger the intrasubject standard deviation in estimation, and the smaller the intrasubject standard deviation in production.

*b. Anxiety and parameters.* Correlation coefficients were calculated between anxiety  $Z$ -scores and intercepts or arctan values representative of clock rates for each block of trials, each method, and each group. Average correlation coefficients, obtained via Fisher's  $r$  to  $Z$  transformation, are listed in Table 8 for each group, each anxiety measure, and each method. It can be seen that whereas average coefficients involving intercepts are variable in sign and while only one of eight approaches significance, all coefficients involving clock rate turned out to be positive, five out of eight being significant at or beyond the .05 level. In consideration of the small numbers of  $S$ s and effects of error, these findings are interpreted as being consonant with Hypothesis 4.2: vis., that anxiety and intercepts are unrelated, while anxiety and clock rates (arctan values representative of clock rates) are positively

TABLE 7  
 MEAN  $r$ 'S (VIA FISHER'S  $r$  TO  $Z$  TRANSFORMATION) ACROSS STANDARDS OF CORRELATIONS BETWEEN  
 ANXIETY  $Z$ -SCORES AND MEAN JUDGMENTS AND BETWEEN ANXIETY  $Z$ -SCORES  
 AND INTRASUBJECT VARIABILITY

Groups	Between anxiety $Z$ -scores and mean judgments (across blocks)			Between anxiety $Z$ -scores and intrasubject variability of judgments		
	Estimation	Production		Estimation	Production	
			<i>SD</i>	<i>CV</i>	<i>SD</i>	<i>CV</i>
Normals						
Taylor MAS	+ .53***	-.53***	+ .31**	-.36**	-.51***	-.25*
MAACL	+ .23**	-.17 (NS)	+ .02 (NS)	-.24**	-.06 (NS)	+ .08 (NS)
Patients						
Taylor MAS	+ .16 (NS)	-.23 (NS)	+ .06 (NS)	-.23 (NS)	-.16 (NS)	+ .05 (NS)
MAACL	+ .37**	-.26**	+ .27**	-.10 (NS)	-.25**	-.12 (NS)

Note: Taylor MAS = Taylor Manifest Anxiety Scale; MAACL = Today Anxiety scale of the Multiple Affect Adjective Check List; CV = coefficients of variation.

\* One-tailed. All other probability values based on two-tailed tests.

\*  $p < .10$ .

\*\*  $p < .05$ .

\*\*\*  $p < .001$ .

correlated. However, the predictions on differences between trait anxiety and state anxiety (Hypothesis 4.3) were not borne out.

In the following, some correlations between anxiety and other statistics derived from time judgments will be presented. Although most of these correlations are not significantly different from zero, their patterns may suggest future experiments.

Correlation coefficients between anxiety and intrasubject variability of parameters are listed in Table 9. The table suggests that the relation between anxiety and standard deviations of intercepts varies with the method, but

TABLE 8  
MEAN  $r$ 'S (VIA FISHER'S  $r$  TO  $Z$  TRANSFORMATION) ACROSS BLOCKS OF TRIALS OF  
CORRELATIONS BETWEEN ANXIETY  $Z$ -SCORES AND INTERCEPTS AND  
ARCTAN VALUES REPRESENTATIVE OF CLOCK RATES

Groups	Intercepts		Arctan values	
	Estimation	Production	Estimation	Production
Normals				
Taylor MAS	-.14 (NS)	+.24 <sup>a**</sup>	+.58 <sup>****</sup>	+.52 <sup>****</sup>
MAACL	-.09 (NS)	+.02 (NS)	+.25 <sup>**</sup>	+.18 <sup>a*</sup>
Patients				
Taylor MAS	-.08 (NS)	-.01 (NS)	+.32 <sup>**</sup>	+.23 <sup>a*</sup>
MAACL	+.02 (NS)	-.17 (NS)	+.41 <sup>***</sup>	+.16 (NS)

Note: Taylor MAS = Taylor Manifest Anxiety Scale; MAACL = Today Anxiety scale of the Multiple Affect Adjective Check List.

<sup>a</sup> One-tailed. All other probability values based on two-tailed tests.

\*  $p < .10$ .

\*\*  $p < .05$ .

\*\*\*  $p < .01$ .

\*\*\*\*  $p < .001$ .

that the relationship between anxiety and variability of arctan values representative of clock rates may be negative.

Correlation coefficients between anxiety and rate of change of parameters per block, and between anxiety and rate of change of parameters per block expressed as a fraction of the respective intrasubject standard deviations, are given in Table 10. Here, coefficients involving rate of change of intercepts appear not to suggest a pattern. Without permitting a definite answer to Hypothesis 4.5, coefficients involving changes in clock rate suggest that among normals, the higher the anxiety, the smaller the rate of slowing of the internal clock, while among patients, the higher the anxiety, the greater the rate of slowing of the internal clock.

TABLE 9  
CORRELATION COEFFICIENTS BETWEEN ANXIETY Z-SCORES AND (INTRASUBJECT) VARIABILITY  
OF INTERCEPTS AND ARCTAN VALUES REPRESENTATIVE OF CLOCK RATES

Groups	Intercepts		Arctan values			
	Estimation	Production	SD	CV	SD	CV
Normals						
Taylor MAS	.57	-.55	-.12	-.40	-.42	-.61*
MAACL	.29	-.10	-.45	-.33	-.17	-.21
Patients						
Taylor MAS	.41	-.77**	-.03	-.14	-.53	-.49
MAACL	.53	-.58	.12	-.05	-.34	-.35

Note: SDs and coefficients of variation (CV) relative to mean arctan value representative of clock rate. Taylor MAS =

Taylor Manifest Anxiety Scale; MAACL = Today Anxiety scale of the Multiple Affect Adjective Check List

\*  $p < .10$  (two-tailed).

\*\*  $p < .05$  (two-tailed).

TABLE 10  
CORRELATION COEFFICIENTS BETWEEN ANXIETY Z-SCORES AND RATE OF CHANGE OF PARAMETERS PER  
BLOCK ("RATE") AND RATE OF CHANGE OF PARAMETERS PER BLOCK RELATIVE TO THE  
RESPECTIVE INTRASUBJECT STANDARD DEVIATION ("REL. RATE")

Groups	Intercepts			Production			Estimation			Clock rate (arctan slopes)		
	Estimation		Rate	Production		Rate	Estimation		Rate	Clock rate (arctan slopes)		Rel. rate
	Rate	Rel. rate		Rate	Rel. rate		Rate	Rel. rate		Rate	Rel. rate	
Normals												
Taylor MAS	.52	.50	.30	.03	.24	.22	.24	.22	.52	.42	.42	.50
MAACL	-.15	-.26	.12	.08	.62*	.35	.62*	.35	.42	.50	.50	.50
Patients												
Taylor MAS	.71*	.61	-.14	-.17	-.33	-.28	-.33	-.28	-.70*	-.75*	-.75*	-.75*
MAACL	.50	.43	-.29	-.40	-.39	-.28	-.39	-.28	-.47	-.54	-.54	-.54

Note: Arctan values representative of clock rates were used instead of slopes. Taylor MAS = Taylor Manifest Anxiety Scale; MAACL = Today Anxiety scale of the Multiple Affect Adjective Check List.

\*  $p < .10$  (two-tailed).



### 6. On Normal-Patient Differences

It appears that the present experiment does not demonstrate major differences between linear normals and linear patients. Because of small numbers of Ss and effects of unreliability, differences between means, between variabilities, and between strengths of association between variables may not have reached statistical significance. Thus, the present experiment falls short of yielding information on possible differences in the degree of positive correlation between anxiety and clock rates, whether between methods or between the two groups of Ss. However, in spite of these limitations, the existence of correlations between anxiety and time judgments in both groups, especially between anxiety and arctan values representative of clock rates, implies that comparisons between different nosological groups may lead to erroneous conclusions unless confounding by possible anxiety differences is avoided.

Differences between groups as found in the present experiment, perhaps to be regarded as minor differences, cannot in any case be directly attributed to specific differences between normals and chronic undifferentiated schizophrenics: not only did groups differ in drug regimen (absence of known medication in normals, presence of variable medication among patients), but contributions from sampling variation cannot be ruled out. These minor differences exist with respect to clock rates and possibly with respect to the relationship between anxiety and the rate of slowing of the internal clock over blocks of trials. As to clock rates, there are no significant differences between methods for normals, whereas in patients, the internal clock ran significantly faster in magnitude production than in magnitude estimation. Concerning anxiety and the rate of slowing, the data suggest that among normals, the higher the anxiety, the smaller the rate of slowing, whereas among patients, the higher the anxiety, the greater the rate of slowing over blocks of trials.

### REFERENCES

1. ADLER, N. Perception of time as a function of self-organization. Unpublished Ph.D. dissertation, University of California, Berkeley, 1954.
2. ANLIKER, J. E. Brain waves and operant behavior. *J. Exper. Anal. Behav.*, 1959, **2**, 252 (abs.).
3. ———. Variations in alpha voltage of the electroencephalogram and time perception. *Science*, 1963, **140**, 1307-1309.
4. BERGLUND, B., BERGLUND, U., EKMAN, G., & FRANKENHAEUSER, M. The influence of auditory stimulus intensity on apparent duration. *Scand J. Psychol.*, 1969, **10**, 21-26.

5. BINDRA, D., & WAKSBERG, H. Methods and terminology in studies of time-estimation. *Psychol. Bull.*, 1956, **53**, 155-159.
6. BRACKBILL, G., & LITTLE, K. MMPI correlates of the Taylor Scale of manifest anxiety. *J. Consult. Psychol.*, 1954, **18**, 433-436.
7. BROWNE, W. A. F. The perception, &c., of time as a feature in mental disease. *Brit. J. Ment. Sci.*, 1874, **19**, 519-532.
8. BURNS, N. M., & GIFFORD, E. C. Time estimation and anxiety. *J. Psychol. Stud.*, 1961, **12**, 19-24.
9. CAHOON, R. L. Physiological arousal and time estimation. *Percept. & Motor Skills*, 1969, **28**, 259-268.
10. CARLSON, V. R., & FEINBERG, I. Individual variations in time judgment and the concept of an internal clock. *J. Exper. Psychol.*, 1968, **77**, 631-640.
11. ———. Time judgment as a function of method, practice, and sex. *J. Exper. Psychol.*, 1970, **85**, 171-180.
12. CLAUSEN, J. An evaluation of experimental methods of time judgment. *J. Exper. Psychol.*, 1950, **40**, 756-761.
13. COHEN, S. I., & MEZEY, A. G. The effect of anxiety on time judgment and time experience in normal persons. *J. Neurol. Neurosurg. & Psychiat.*, 1961, **24**, 266-268.
14. DOBSON, W. R. An investigation of various factors involved in time perception as manifested by different nosological groups. *J. Gen. Psychol.*, 1954, **50**, 277-298.
15. DUREMAN, I., & EDSTRÖM, R. EEG and time perception. Report No. 22, Department of Psychology, University of Uppsala, Sweden, December 1964.
16. DUREMAN, I., & PALSHAMMAR, S.-A. Psychophysiological reactions in a serial approach-avoidance conflict situation: Its relation to test anxiety and school motivation in children. Report No. 51, Department of Psychology, University of Uppsala, Sweden, April 1968.
17. FALK, J. L., & BINDRA, D. Judgment of time as a function of serial position and stress. *J. Exper. Psychol.*, 1954, **47**, 279-282.
18. FRAISSE, P. Etude comparée de la perception et de l'estimation de la durée chez les enfants et les adultes. *Enfance*, 1948, **1**, 199-211.
19. FRANCOIS, M., Contribution à l'étude du sens du temps: La température interne comme facteur de variation de l'appréciation subjective des durées. *Ann. Psychol.*, 1927, **28**, 186-204.
20. FRANKENHAEUSER, M. Estimation of Time, An Experimental Study. Stockholm, Sweden: Almqvist & Wiksell, 1959.
21. GOLDSTEIN, L., & SUGERMAN, A. A. EEG correlates of psychopathology. Chap. 1 in J. Zubin & C. Shagass (Eds.), *Neurobiological Aspects of Psychopathology*. New York: Grune & Stratton, 1969.
22. GOLDSTONE, S., & GOLDFARB, J. L. Time estimation and psychopathology. *Percept. & Motor Skills*, 1962, **15**, 28.
23. GOLDSTONE, S., BOARDMAN, W. K., & LHAMON, W. T. Kinesthetic cues in the development of time concepts. *J. Genet. Psychol.*, 1958, **93**, 185-190.
24. ———. Effect of quinal barbitone, dextro-amphetamine, and placebo on apparent time. *Brit. J. Psychol.*, 1958, **49**, 324-328.
25. HARE, R. D. The estimation of short temporal intervals terminated by shock. *J. Clin. Psychol.*, 1963, **19**, 378-380.
26. ———. Anxiety, temporal estimation, and rate of counting. *Percept. & Motor Skills*, 1963, **16**, 441-444.
27. HOAGLAND, H. The physiological control of judgments of duration: Evidence for a chemical clock. *J. Gen. Psychol.*, 1933, **9**, 267-287.
28. LACEY, J. I. Somatic response patterning and stress: Some revisions of activation theory. In M. H. Appley & R. Trumbull (Eds.), *Psychological Stress*. New York: Appleton, 1967.

29. LANGER, J., WAPNER, S., & WERNER, H. The effects of danger upon the experience of time. *Amer. J. Psychol.*, 1961, **74**, 94-97.
30. LEGG, C. F. Alpha rhythm and time judgments. *J. Exper. Psychol.*, 1968, **78**, 46-49.
31. LHAMON, W. T., & GOLDSTONE, S. The time sense: Estimation of one second durations by schizophrenic patients. *Arch. Neurol. & Psychiat.*, 1956, **76**, 625-629.
32. MCCONCHIE, R. D., & RUTSCHMANN, J. Human time estimation: On differences between methods. *Percept. & Motor Skills*, 1971, **32**, 319-336.
33. MEZEY, A. G., & COHEN, S. I. The effect of depressive illness on time judgment and time experience. *J. Neurol. Neurosurg. & Psychiat.*, 1961, **24**, 269-270.
34. PEARL, D., & BERG, P. S. Time perception and conflict arousal in schizophrenia. *J. Abn. & Soc. Psychol.*, 1963, **66**, 332-338.
35. PIÉRON, H. *The Sensations: Their Functions, Processes and Mechanisms*. New Haven: Yale Univ. Press, 1952.
36. SAMPSON, H., & BINDRA, D. "Manifest" anxiety, neurotic anxiety, and the rate of conditioning. *J. Abn. & Soc. Psychol.*, 1954, **49**, 256-259.
37. SCHILDER, P. Psychopathology of time. *J. Nerv. & Ment. Dis.*, 1936, **83**, 530-546.
38. ———. *Mind: Perception and Thought in Their Constructive Aspects*. New York: Columbia Univ. Press, 1942.
39. SIEGMAN, A. W. Anxiety, impulse control, intelligence, and the estimation of time. *J. Clin. Psychol.*, 1962, **18**, 103-105.
40. SPREEN, O. The position of time estimation in a factor analysis and its relation to some personality variables. *Psychol. Rec.*, 1963, **13**, 455-464.
41. TREISMAN, M. Temporal discrimination and the indifference interval: Implications for a model of the "internal clock." *Psychol. Monog.*, 1963, **77**(13), Whole No. 576.
42. VENABLES, P. H., & WING, J. K. Level of arousal and the sub-classification of schizophrenia. *Arch. Gen. Psychiat.*, 1962, **7**, 114-119.
43. WARM, J., MORRIS, J., & KEW, J. Temporal judgment as a function of nosological classification and experimental method. *J. of Psychol.*, 1963, **55**, 287-297.
44. WEBSTER, F. R., GOLDSTONE, S., & WEBB, W. W. Time judgment and schizophrenia: Psychophysical method as a relevant contextual factor. *J. of Psychol.*, 1962, **54**, 159-164.
45. WEINSTEIN, A. D., GOLDSTONE, S., & BOARDMAN, W. K. The effect of recent and remote frames of reference on temporal judgments of schizophrenic patients. *J. Abn. & Soc. Psychol.*, 1958, **57**, 241-244.
46. WENGER, M. A. Studies of autonomic balance: A summary. *Psychophysiol.*, 1966, **2**, 173-186.
47. WERBOFF, J. The relation between electroencephalographic activity and the estimation of short temporal intervals. *Diss. Abst.*, 1957, **17**, 2325.
48. ———. Time judgment as a function of electroencephalographic activity. *Exper. Neurol.*, 1962, **6**, 152-160.
49. WRIGHT, D. J., GOLDSTONE, S., & BOARDMAN, W. K. Time judgment and schizophrenia: Step interval as a relevant contextual factor. *J. of Psychol.*, 1962, **54**, 33-38.
50. ZUCKERMAN, M., & LUBIN, B. *Manual for the Multiple Affect Adjective Check List*. San Diego: Educ. & Indus. Test. Serv., 1965.

Department of Research Psychology  
 New York State Psychiatric Institute  
 722 West 168 Street  
 New York, New York 10032



## A PRELIMINARY NOTE ON DEMOGRAPHIC AND PERSONALITY CORRELATES OF DECUBITUS ULCER INCIDENCE\*

*Virginia Commonwealth University and Woodrow Wilson Rehabilitation Center,  
Fishersville, Virginia*

---

JOHN G. CULL AND OTHO H. SMITH

---

### SUMMARY

The purpose of this study was to investigate the correlation between the incidence of decubitus ulcers and the level of spinal cord injury, sex, race, level of intellectual functioning, and level of academic achievement. A total of 114 subjects with spinal cord injuries were surveyed and tested at the Woodrow Wilson Rehabilitation Center, Fishersville, Virginia. The results of this study indicated no relationship between the incidence of decubitus ulcers and the demographic or personality factors studied.

---

### A. INTRODUCTION

Decubitus ulcers are generally acknowledged to be a major stumbling block in the total rehabilitation of the paraplegic and quadriplegic individual. While ulcers can be quite costly in terms of funds, as well as staff time, they oftentimes render the vocational adjustment of the individual impossible.

There are numerous references in the literature relative to the physical etiological factors of decubitus ulcers (pressure, heat, moisture, shearing forces, friction, nutrition, and the general hygiene of the patient); however, there is a striking paucity of references relating to the psychological and demographic factors that predispose the individual to an increased probability of ulceration. Current etiological studies in the literature have little or no predictive value. The purpose of this study is to identify those psychological and demographic factors that have a high degree of correlation with the incidence of decubitus ulcers.

Pumphrey (2) felt in her study of decubitus ulcer patients there were overriding but undetermined personality factors that played a selective role in the incidence of decubitus ulcers. This current study is to investigate the correlation between the incidence of decubitus ulcers and the level of spinal

---

\* Received in the Editorial Office on July 25, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.



cord injury, sex, race, level of intellectual functioning, and level of academic achievement.

It was hypothesized that (a) the higher the level of spinal cord injury the higher the incidence of decubitus ulcers in the cord injured; (b) sex is a selective factor in that females have fewer ulcers than males; (c) nonwhites have fewer decubitus ulcers than whites; (d) the higher the level of intellectual functioning, the lower the incidence of decubitus ulcers; (e) the higher the academic achievement, the fewer ulcers the cord injured will have.

### B. METHODOLOGY

For the purposes of this study, a total of 114 cord injured students who were engaged (or recently had been engaged) in a vocational rehabilitation program at the Woodrow Wilson Rehabilitation Center, Fishersville, Virginia, were surveyed and tested. The Woodrow Wilson Rehabilitation Center is a comprehensive, vocationally oriented rehabilitation center with a medical service. These students were identified as to level of cord injury, sex, race, and age.

The level of intellectual functioning was determined by the administration of the Wechsler Adult Intelligence Scale, and the level of academic functioning was assessed by the use of the Wide Range Achievement Test (1). A survey form was completed by each subject. These data then were analyzed statistically by the appropriate statistical tests as indicated below.

### C. RESULTS

None of the hypotheses was strongly supported. The differences between the mean number of decubitus ulcers for the quadriplegics and paraplegics yielded a calculated  $t$  of .83 with 111 degrees of freedom, which is not significant at the .05 level. The level of injury in the incidence of ulcers in the students again was not statistically significant (calculations yielded a chi square of .002 with one degree of freedom). This same condition was obtained when investigating the differences between race and the incidence of ulcer occurrence (chi square of .008 with one degree of freedom). However, there was a difference on the sex factor at the .05 level (calculated chi square of 2.12 with one degree of freedom).

On the intelligence testing and on the various scales of the Wide Range Achievement Test, there were no statistically significant discriminating factors that could be used to predict the incidence of decubitus ulcers. Therefore, the hypotheses relating to intelligence and academic achievement were not supported.

## D. DISCUSSION

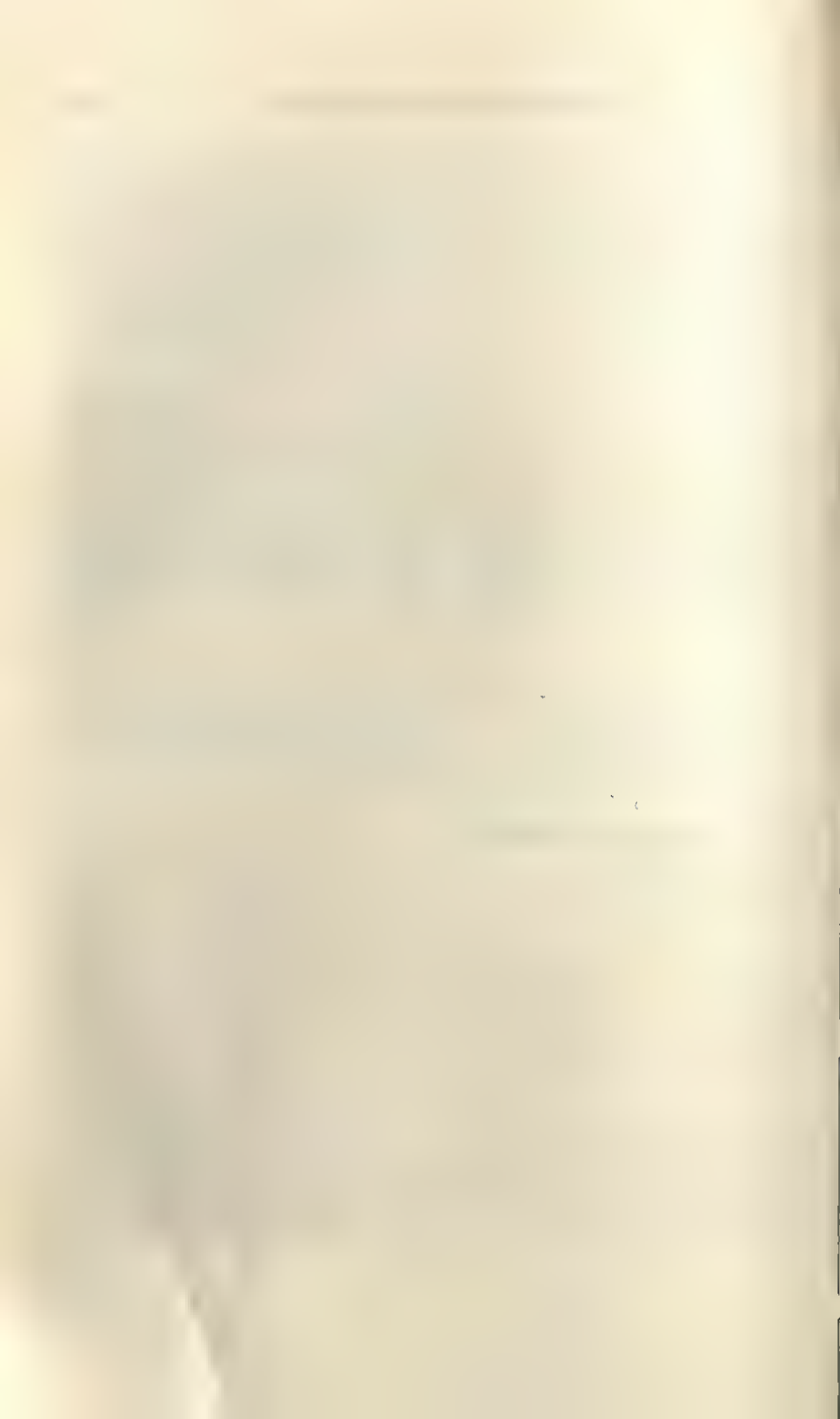
The data gathered on the 114 cord injured students undergoing rehabilitation services indicate that the hypotheses are not tenable. The data indicate that sex is a slight selective factor in that females have fewer ulcers than males; however, it is felt that sex is not a predictive demographic factor in the incidence of decubitus ulcers. The reason for this slight selective sex factor is due to physiological conditions: namely, the heightened concentration of fatty tissue which functions as a cushion in the female.

Since none of the demographic or personality factors considered in this study proved to have a predictive value for decubitus ulcer incidence, we feel research should continue in this area to determine selective factors that will assist in predicting decubitus ulcer incidence. This is a key area of research in that decubitus ulcers can be devastating to the smooth progress of rehabilitation. The incidence of ulcers disrupts the students as well as the process. Therefore, we feel that it is urgent that some predictive indices be developed to anticipate the decubitus ulcer-prone paraplegic or quadriplegic. This study failed to uncover any such indices.

## REFERENCES

1. JASTAK, J. F., & JASTAK, S. R. The Wide Range Achievement Test Manual. Wilmington, Del.: Guid. Assoc., 1965.
2. PUMPHREY, K. The pressure sore: A complication of paraplegia. Unpublished manuscript, University of Virginia, Charlottesville, 1966.

*Regional Counselor Training Program*  
*Virginia Commonwealth University*  
*P. O. Box 499 WWRC*  
*Fishersville, Virginia 22939*



## THE RELATION OF STUDENT AND TEACHER TRAITS OF AUTHORITARIANISM TO STUDENT ACHIEVEMENT IN ENGLISH\*

*University of South Carolina*

---

JOHN G. WAY

---

### SUMMARY

The study investigated the relationship of five selected traits of authoritarianism, possessed by both college freshmen and their instructors, to student achievement in freshman English when student verbal aptitude was not a factor. The results indicated that student achievement level was affected more by instructor traits than by traits of the students themselves, although student traits did serve to qualify the effect of instructor traits.

---

### A. INTRODUCTION

Investigators have been interested in the effect nonintellectual variables have on student achievement at the college level for over half a century. Most of their research has dealt with either teacher personality or student personality. Only a few of the reported studies have been designed to discover the interrelations of selected student and teacher personality traits that affect student scholastic performance (1, 4, 5).

Although Milholland (5) reported encouraging results from matching students with instructors on the basis of personality trait interactions that had indicated significant relations to course outcome, there still remains a paucity of research regarding such interactions and interrelation effects.

The present study was designed to investigate whether or not five selected traits of authoritarianism, possessed in varying degrees by both college freshmen and their instructors, influence student achievement in freshmen English when student verbal aptitude is not a factor.

### B. METHOD

#### 1. Subjects

Twenty-two male and 10 female graduate students in the English Department at the University of North Carolina at Chapel Hill (UNC), serving

---

\* Received in the Editorial Office on July 25, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

as instructors of freshman English, and 508 male students enrolled in their sections during the fall of 1967 were selected as the pools from which the instructor and student samples were drawn because (a) the number of available instructors was relatively large, (b) the differences in age and teaching experience of the instructors were minimized (mean age of 27 years), (c) the number of students in each section was approximately the same (maximum number of 22), (d) the course has a common syllabus and is required of all freshman students, and (e) since verbal scores on the *College Entrance Examination Board* (CEEB) were the best available index of verbal ability for purposes of the covariance variable, it was felt that these scores would be more directly related to achievement in the freshman English course content (composition and rhetoric) than they might be for other courses. The student sample included only male students, since UNC typically has only a nominal enrollment of females in the freshman class.

## 2. Procedure

The *Pensacola Z Survey* (PZS) is an instrument developed by Jones (3) to be a comprehensive measure of the core traits constituting the authoritarian syndrome. It is relatively unsusceptible to faking and free from specific subcultural effects inherent in other instruments purported to measure the same traits. The traits measured by the PZS are dependency, hostility, anxiety, rigidity, and heteronomy (a combination of the first four traits).

The PZS was administered to both the instructor and student samples during a two week period near the end of the fall semester. Each distribution of trait scores for the instructor sample and the student sample was divided into approximately equal high, medium, and low trait groups. They were further divided on the basis of instructor sex.

Final course grade was used as the index of student achievement. Instructors assigned their students qualified letter grades which were converted to a numerical grade ranging from a high of 12 to a low of zero. Although an objective common course examination was not used to assign grades, each instructor followed a common syllabus and used a common guide for assigning coursework and evaluating overall performance in the course.

In order to adjust for differences in verbal ability among the student participants, CEEB verbal scores were obtained from a roster of such information provided by the UNC office of central records.



### C. RESULTS

Analyses of covariance were performed on the converted final course grades of students who had been subgrouped on the basis of student personality traits, instructor personality traits, and instructor sex. A separate analysis was performed for the student subgroups derived from student and instructor rankings on each of the five PZS scales.

#### 1. *Main Effects*

The analyses failed to reveal any differences among the achievement levels due to the main effects of student personality traits or instructor's sex. Significant differences in level of achievement did occur among students who had received instruction from combined male and female instructors ranked high, medium, or low on the Anxiety Scale ( $F = 4.288$ ;  $p < .05$ ), Hostility Scale ( $F = 8.549$ ;  $p < .001$ ), and Heteronomy Scale ( $F = 5.104$ ;  $p < .001$ ). These main effects were qualified by the following simple effects (6).

#### 2. *Simple Effects*

a. *Instructor personality and instructor sex.* Differences in mean achievement levels occurred among students in sections taught by male instructors variously ranked on the Rigidity Scale ( $F = 6.208$ ;  $p < .001$ ), Hostility Scale ( $F = 4.451$ ;  $p < .01$ ), and Anxiety Scale ( $F = 5.880$ ;  $p < .01$ ). Differences occurred among the achievement levels of students in sections taught by female instructors ranked in terms of their scores on the Hostility Scale ( $F = 3.340$ ;  $p < .05$ ).

b. *Student personality and instructor sex.* No differences were detected among the achievement levels of students when student personality and instructor sex were the variables from which the student subgroups were derived.

c. *Student personality and instructor personality.* Differences in mean achievement levels occurred only when student subgroups were derived from student and teacher rankings on the Hostility Scale ( $F = 1.984$ ;  $p < .05$ ).

d. *Student personality, instructor personality, and instructor sex.* It was found that differences in achievement levels occurred among student subgroups in sections taught by male instructors when the subgroups were derived from student and instructor rankings on the Rigidity Scale ( $F = 2.474$ ;  $p < .05$ ) and Anxiety Scale ( $F = 1.963$ ;  $p < .05$ ). The only significant difference in achievement levels of students in sections taught by

female instructors was among the subgroups derived from student and teacher rankings on the Hostility Scale ( $F = 2.170$ ;  $p < .05$ ).

#### D. DISCUSSION

In general, the findings of this study indicated that student achievement level tended to be more affected by teacher traits than by traits of the students themselves, although student traits qualified the effect of instructor traits. Further analyses, using Scheffé's method, revealed that a positive relationship existed between student achievement level and the instructor traits of rigidity and anxiety. If it could be assumed that teachers who were ranked high on these traits employed more autocratic teaching methods than did those who were ranked lower, these findings would lend support to those of Guetzkow, Kelly, and McKeachie (2) who found that students achieved at a higher level in classes taught by autocratic methods as compared to their achievement level when less autocratic methods were used. In the present study, however, it was impossible to know whether the students in sections taught by instructors ranked high on the authoritarian traits actually achieved at a higher level or whether an adjunct characteristic of teachers ranked high on such traits is that they also tend to assign higher grades than do instructors ranked relatively lower.

It is interesting to note that of the authoritarian traits studied, anxiety, rigidity, and hostility appeared to be exerting the greater influence on student achievement. The scales measuring these traits were the ones that had been included in the PZS on the basis of having been empirically derived (3). Dependency, the trait that had been arbitrarily included because it logically represented the most basic trait of the authoritarian syndrome, failed to reveal an influence strong enough to be detected. It is possible that a more sensitive instrument for measuring dependency would have revealed that it also exerts an influence on student achievement.

Although instructor sex differences did not by themselves appear to influence student achievement, they did qualify the main effect of teacher personality on achievement. Generally, it appeared that the personality traits of male instructors influenced student grades more than did such traits of female instructors. When instructor sex was ignored, instructor rigidity did not significantly differentiate among levels of student achievement. However, when sex was considered, instructor rigidity detected a highly significant difference ( $p < .001$ ) among students taught by males. In addition, although instructor hostility was found to differentiate among levels of

achievement for both male and female instructors, the difference was more pronounced in the case of males. Heteronomy, the conglomerate trait, detected a highly significant difference among achievement levels ( $p < .001$ ) only when sex differences were ignored.

Further analysis of the simple effects of student personality, teacher personality, and teacher sex revealed that students ranked high or low on rigidity, anxiety, and hostility paired with instructors also ranked high or low on these same traits achieved at a higher level than did any group of students paired with instructors ranked medium. Attempts to interpret this finding would be merely speculative, since it is not known whether or not grades represent actual achievement.

The importance of the findings of this study is not so much that of demonstrating the relationship of specific personality traits to student achievement as it is in demonstrating the qualifying effect that student and teacher traits, when taken together, have on one another in terms of accounting for variation in student achievement. Although the simple effects reported in the study are not synonymous with interaction effects, they do strongly suggest such interactions. In addition, it appears that teachers should be cautious in assigning grades because such grades may not validly represent achievement. Society, however, reacts to the grade as if it were an index of actual academic achievement and potential.

Most of the past studies relating various personality variables to student achievement resulted in apparently inconsistent or contradictory findings. Except for a very few, the studies considered only the dimension of student traits or instructor traits. The findings of these studies might have been more consistent if the two dimensions had been studied in a factorial design.

#### REFERENCES

1. BAERWALDT, J. W., MILHOLLAND, J. E., & McKEACHIE, W. J. Student ratings of teachers and student performance as a function of student and teacher personality factors. In W. J. McKeachie, R. L. Isaacson, & J. E. Milholland (Eds.), *The Michigan Project: Research on Characteristics of Effective College Teaching*. Ann Arbor: Univ. Michigan, 1963.
2. GUSTEKOW, H., KELLY, E. L., & McKEACHIE, W. J. An experimental comparison of recitation, discussion and tutorial methods in college teaching. *Educ. Psychol.*, 1954, 45, 193-207.
3. JONES, M. B. The Pensacola Z Survey: A study in the measurement of authoritarian tendency. *Psychol. Monog.*, 1957, 71(23), Whole No. 452.
4. McKEACHIE, W. J. Research on teaching at the college and university level. In N. L. Gage (Ed.), *Handbook of Research on Teaching*. Chicago: Rand McNally, 1963.
5. MILHOLLAND, J. E. The matching of students and teachers so as to enhance student achievement. In W. J. McKeachie, R. L. Isaacson, & J. E. Milholland (Eds.),

*The Michigan Project: Research on Characteristics of Effective College Teaching.*  
Ann Arbor: Univ. Michigan, 1963.

6. WINER, B. J. *Statistical Principles in Experimental Design.* New York: McGraw-Hill, 1962.

746 Della Drive

Lexington, Kentucky 40504

## THE STRUCTURE OF BEHAVIORAL VALUES OF COLLEGE STUDENTS\*

*Michigan State University*

---

JOHN PAUL MCKINNEY<sup>1</sup>

---

### SUMMARY

A factor analysis of 744 students' social desirability ratings for 95 questionnaire items yielded nine factors of behavioral values, eight of which were considered reliable. The questionnaire was comprised of the universe of responses given by an earlier sample of subjects to the incomplete sentence stems, "I would be ashamed of myself if . . ." and "I am proud of myself when . . . ." Thus, the items were not preselected on any theoretical basis, but were established empirically as the stated behavioral values of a group of college students. The resulting factors suggested an organization of values that distinguishes between competence and moral values, between personal and interpersonal values, and between subject and object values.

---

### A. INTRODUCTION

A number of investigators studying the behavioral values of youth have attempted to discover the underlying value dimensions that are relevant for this age group. Rettig and Pasamanick (12), for example, factor analyzed the moral judgment ratings of a group of college students in 1958 on a set of items that had also been given to similar groups in 1929, 1939, and 1949 (5, 6). They found that the Crissman questionnaire yielded the following factors: basic morality (a general factor); religious moral value; family maintenance; puritanical morality; predelinquent morality; and economic morality.

Another factor analytic approach was used by Brim (4) and Goodman

---

\* Received in the Editorial Office on July 26, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The author would like to thank Professor Eugene Jacobson, who first suggested this analysis and who has continued to provide many constructive comments about the research. He is grateful, too, to Miss Joyce Pennington, Mrs. Margaret Connolly, and Mrs. Esther Cohen, for aid in data collection and analysis, and to Dr. Stanley Cohen, for assistance in computer programming.



(8) in the analysis of a set of prescriptive behavioral statements designed to discover the value structure of a group of high school students. Basically, the researchers found that, for items emphasizing achievement, adolescents' values were organized around three important constructs: (a) the source of the value (i. e., self or other); (b) the various roles that the respondent might assume (e. g., student, peer group member, family member, etc.); and (c) the theme or content of the behaviors, (e. g., achievement).

By factor analyzing the results of questionnaire data given to Russian and English boys and girls, Rodgers, Bronfenbrenner, and Devereux (13) found that for both groups, two factors, which they called "manners" and "masculinity," emerged. As the authors noted, these two factors bear a striking similarity to two of the meaning dimensions described by Osgood (10): namely, evaluation and potency. In other words, here is evidence that these meaning dimensions can describe not only the cognitive organization of a wide variety of objects, but also the organization of behavioral standards.

In each of these studies the items used for the factor analysis were selected by the researchers for a specific purpose. An alternative approach would be an open-ended procedure by which the items could be produced by the respondents themselves. With this purely empirical approach, it would be possible to use the universe of value statements given by a sample of subjects. Thus one could study the subjective organization of such values as given by the subjects, rather than as based on a predetermined questionnaire, which is necessarily limited by the hypotheses of the researcher.

The purpose of the present study was to uncover the prescriptive behavioral values of college students and to describe the subjective organization of these values.

## B. METHOD

The procedure involved the use of a 95 item questionnaire which was developed from an earlier instrument. The earlier test was a sentence completion form, used in a study of the development of values (9). The sentence completion test consisted of 20 stems. One of those stems contained a statement of pride and another a statement of shame: "I would be ashamed of myself if . . ." and "I am proud of myself when . . ."

The sentence completion test had been administered to 67 college students. The universe of their responses to these two statements of pride and shame constituted the questionnaire used in the present study. The list consisted of 95 items after the exact repetitions were discarded. Any item that differed from the others on the list, however slight that difference may have been,

was retained. Thus the final questionnaire included as different items, such similar statements as "getting good grades," "making a good grade," and "getting a good grade for a class." It was felt it would be better to err in the direction of overinclusion than to leave out an item that could contain additional meaning through a subtle difference in wording.

The 95 item questionnaire was then given as a Likert-type scale to a new group of Ss, 744 university students enrolled in an introductory psychology course, with the following instructions:

Below you will find examples of things a person says he does or does not do. On the IBM scoring sheet (next to the item that corresponds to the behavior) mark the number that reflects how socially desirable or socially undesirable you think the behavior is: 1. Desirable; 2. Moderately desirable; 3. Neutral, neither desirable nor undesirable; 4. Moderately undesirable; 5. Undesirable. You are to judge the behavior in terms of whether you consider it to be desirable or undesirable in others. Be sure to make a statement about each behavior.

Their responses were then subjected to a principal components factor analysis, and the factors were rotated to a varimax solution.

### C. RESULTS

The resulting factor solution yielded nine factors which together accounted for approximately 43 percent of the total variance. Table 1 presents the rotated factors with the loadings for each item. All items with loadings of .30 and higher have been included (except for factors A, B, and C for which only items with loadings over .35 have been included).

The reliability of the factors was determined by a split-half technique described by Barlow and Burt (2). Basically, the technique involves splitting the subjects into two equal groups and obtaining separate factor analyses from each group. The two analyses are then compared with use of a "coefficient of similarity" ( $s$ ), a correlational technique. These split-half correlations ( $s$ ) are given in Table 1 for each factor. Six of the factors had acceptable reliability coefficients above .70. Two of the remaining three factors had marginally acceptable reliabilities of .58 and .66, and one factor had an unacceptable reliability of .27.

### D. DISCUSSION

Factor A, which accounts for the largest percentage of the total variance (11%) is clearly an academic achievement factor. It should come as no surprise, when one considers the vocation of the subjects, that most of their

TABLE 1  
FACTORS OF BEHAVIORAL VALUE STATEMENTS

Item number	Loading	Item description	% <sup>a</sup>	\$ <sup>b</sup>
Factor A			.1070	.98
74	.83	Making a good grade		
70	.82	Getting a good grade for a class		
82	.82	Getting good grades		
60	.81	Doing well in class		
22	.77	Doing well in school		
8	.75	Doing well on a paper or exam		
79	.68	Finishing one's homework completely		
63	.65	Finishing assignments on time		
6	.64	Achieving in school or sports		
11	.62	Receiving a high exam grade through one's own effort		
88	.53	Getting better grades than others		
58	.47	Getting a passing grade in a college class		
68	.45	Succeeding at doing something		
66	.41	Getting dressed up		
49	.38	Doing well in sports		
44	.36	Being well accepted		
45	.36	Accomplishing a goal		
61	.35	Finding oneself leading		
93	.35	Doing well on something one has worked on		
95	-.65	Flunking a test		
13	-.48	Neglecting one's schoolwork		
19	-.45	Lacking a grade because of laziness		
Factor B			.0879	.98
65	-.64	Doing something deceitful		
17	-.62	Lying to someone		
43	-.59	Leading a boy to believe one likes him when one doesn't		
41	-.59	Being jealous of others		
50	-.56	Cutting other people down		
69	-.55	Doing something one thinks is wrong		
39	-.55	Jumping on others for no justified reason		
27	-.52	Ignoring one's friends		
18	-.51	Thinking wrong things of others		
33	-.51	Trying to be something one isn't		
59	-.51	Disliking someone before one knows him		
29	-.49	Letting one's pride get the best of one's self		
94	-.45	Saying something one doesn't mean		
85	-.44	Being two-faced		
46	-.44	Being out of character		
25	-.44	Being irritable		
92	-.43	Indulging in something one promised not to		
56	-.41	Hurting others who love one		
31	-.41	Getting in ugly moods and taking them out on those around one		

TABLE 1 (continued)

Item number	Loading	Item description	% <sup>a</sup>	S <sup>b</sup>
16	— .40	Forgetting about someone's feelings		
28	— .39	Letting someone down		
9	— .36	Getting drunk		
51	— .36	Disappointing those one loves		
84	— .36	Seeming to be phony to one's self		
89	— .35	Failing to honor a commitment		
38	.35	Helping someone out		
48	.35	Helping someone by going out of one's way		
Factor C			.0549	.91
26	.61	Accomplishing something		
45	.61	Accomplishing a goal		
32	.59	Accomplishing something worthwhile		
68	.59	Succeeding at doing something		
7	.57	Doing something well		
67	.54	Setting one's own goals and striving to reach them		
35	.52	Accomplishing something that's very important to one		
14	.49	Accomplishing a difficult task		
93	.45	Doing well on something one has worked on		
36	.35	Doing something creative		
Factor D			.0373	.72
84	.59	Seeming to be phony to one's self		
89	.49	Failing to honor a commitment		
80	.48	Being purposely mean to one's family		
85	.47	Being two-faced		
81	.43	Letting one's self down		
87	.43	Hurting someone's feelings unknowingly		
56	.36	Hurting others who love you		
77	.35	Fighting with one's parents		
94	.35	Saying something one doesn't mean		
86	— .35	Doing something one thinks is good or well done		
Factor E			.0296	.76
44	.60	Being well accepted		
10	.58	Making lots of friends		
37	.57	Being asked out		
49	.41	Doing well in sports		
1	.40	Making people like one		
61	.30	Finding oneself leading		
25	— .33	Being irritable		
Factor F			.0294	.66
55	.57	Doing something really assinine		
72	.56	Making an ass out of one's self at a party		

TABLE 1 (continued)

Item number	Loading	Item description	% <sup>a</sup>	S <sup>b</sup>
90	.56	Drinking too much and making a scene		
73	.46	Making ridiculous errors		
9	.42	Getting drunk		
34	.35	Letting people know one too well		
71	.35	Thinking of some of the things one gets one's self into		
Factor G			.0272	.27
40	.55	Coming up with a new or profound thought		
36	.50	Doing something creative		
78	.47	Doing things one thinks one cannot do		
38	.34	Helping someone out		
52	-.36	Not calling one's parents		
91	-.30	Forgetting to do anything		
Factor H			.0265	.58
3	-.70	Making someone happy		
5	-.37	Doing something constructive		
23	-.34	Helping someone by talking and understanding		
38	-.31	Helping someone out		
4	.65	Hurting someone		
39	.30	Jumping on others for no justified reason		
Factor I			.0211	.76
21	.64	Watching one's diet		
15	.36	Wishing one's friends would go away for a while and leave one alone		
64	.36	Saying no		
83	.30	Refusing food		
24	-.53	Not being able to stick to one's diet		
62	-.44	Overeating		

<sup>a</sup> Percentage of total variance accounted for.

<sup>b</sup> Reliability coefficient.

value statements are organized around a dimension of scholastic performance. There are some rather interesting aspects to their description of academic success, however. First of all, the items are highly specific. Compared to Factor C, also a competence factor, Factor A refers specifically to school performance, while Factor C refers to competence in general. However, these two factors are distinguishable by at least two other characteristics of their respective items.

The items defining Factor A almost all emphasize the *products* of achievement—"a passing grade," "finding oneself leading," "finishing homework,"



etc.—and therefore stress the end result, mainly the evaluation. Secondly, the evaluation is of the finished product, not the quality of the achievement experience.

Factor C, on the other hand, emphasizes general competence, the value of accomplishment. The main distinction from Factor A, aside from C's more general quality, is that Factor C emphasizes the *experience* of achievement rather than the goal. Any evaluation of this experience is clearly internal: e.g., "*setting one's own goals* and striving to reach them," "doing well on something *one has worked on*," "accomplishing something that's *very important to one*." The satisfaction is intrinsic, whereas in Factor A the value was on the external proof of accomplishment: the grade, the exam, the homework. The emphasis in Factor C is rather on the quality of the experience—"doing something *well*," "accomplishing a *difficult task*," "doing something *creative*," "accomplishing something *worthwhile*"—with the criteria for such quality being set intrinsically: ". . . one's own goals," ". . . very important to one," ". . . something one has worked on."

Factor B, which accounted for the second largest percentage of the variance, clearly deals with interpersonal morality. It is distinguished from Factors A and C in a number of interesting ways. Factors A and C deal with competence behaviors, and B deals with moral or ethical concerns. In addition, A and C are exclusively concerned with the individual and his personal achievements, while the items of Factor B deal for the most part with interpersonal events. In other words, the competence values appear to deal with one's self-concerns, while moral values appear to deal with social concerns. Such an organization of values reflects one of the basic classifications given by philosophers in this area. Rescher (11, pp. 17-18), for example, gives as one of the major ways of classifying values, "the relationship between subscriber and beneficiary." About this distinction he says as follows:

In general, a person subscribes to a value because he sees its realization as beneficial to certain people. Consequently yet another approach to the classification of values takes its departure from this point and classifies values according to the "orientation" of the value, that is, according to the relationship that obtains between the person who holds the value, the subscriber, on the one hand, and on the other, the presumptive beneficiaries who benefit from the realization of the value.

This approach leads to a classification of the following sort:

- I. Self-oriented (or egocentric) values . . .
- II. Other-oriented (or disinterested) values . . .

Nor is it surprising that, in our factor analysis, the ethical or moral factor

would be the interpersonal or "other-oriented" factor. Traditionally moral values have dealt with interpersonal relations, and the inclusion of such variables as competence in a values context is relatively recent, according to Aronfreed (1, pp. 2-3):

Moral philosophy generally had taken the view that morality was concerned only with fundamental human relationships. Values were considered to be moral when they were addressed to desires and constraints which had direct implications for the welfare of others. . . . [Whereas] it is quite common today to find that the term conscience is used in a sense that embraces dimensions of value which pertain to cleanliness, to sexual habits, to *proficiency and persistence in achievement*, and to a great many other highly individualized and personal segments of human behavior.<sup>2</sup>

It would appear that the students responding to our questionnaire have maintained this distinction in the organization of their behavioral values.

Interestingly, the first three factors are similar in some respects to a distinction made by Bass (3). Bass observes that in the development of a social group there are three main motivational orientations. Some individuals are motivated by the group itself (that is, the desire for interaction with others), others are motivated by the task, and finally some are motivated by personal concerns. Bass has developed an inventory to measure these three orientations and has reported a number of behavioral and test correlates in a wide variety of situations. Our Factor B appears to be similar to Bass's "interaction-orientation" which he defines in terms of social relations. Our Factor A appears to be related to his "self-orientation," which he describes as characteristic of the person who is "more concerned about his own needs than those of others . . . more interested in extrinsic satisfaction of work . . ." (p. 262). Finally, our Factor C, general competence, appears to be similar to Bass's "task-orientation" which he describes as characteristic of those who try to do their best on assignments. The very items he uses as examples demonstrate this similarity: "To have the feeling of a job well done; . . . To be wise . . . To work at a hobby . . . To be a leader who gets things done" (p. 262).

In addition, Bass reports on a study by Friedlander (7) in which through the factor analysis of questionnaire data he arrived at three main factors contributing to job satisfaction: "recognition through advancement, social environment, and intrinsic self-actualizing work" (3, p. 261). Again these appear to be the same factors reported by our students in describing their general behavioral values. In other words, the organization of students'

<sup>2</sup> Italics added by present author.

values, the behaviors they feel proud or ashamed of, are similar to the individual orientations that have been found to interact significantly with social behavior.

The remaining five factors appear to be more specific in scope. Factor D, which we have labelled "bad faith," deals with hypocrisy, or phoniness. Again, when we consider the student population who served as subjects, it comes as no surprise that self-deceit should be a main concern. The search of youth for authenticity in the establishment of their identities is a well-known theme. It has provided a rich basis for both theorizing about youth and for inspiring a genre of literature that has become increasingly popular. Unsurprisingly then, here is empirical evidence that, indeed, self-consistency is a basic value for youth.

Factor E, "popularity," reflects the subjects' concern with being well-liked. Factor F, "social uncouthness," deals with social unacceptability: namely, those inappropriate social behaviors that would prove embarrassing. Both of these factors, in contrast with the first four, deal with the respondent as object: that is, on the one hand, with being well-liked, and on the other, with being seen as a buffoon.

Factor G, although defined by only a few items and too unreliable to consider anything permanent, seems to refer to "creativity." Factor H defines a concern for "nurturance," or altruism. The final factor "I" refers to self-control, especially with respect to dieting. Four of the six items refer to the highly specific concern of overeating.

In summary, it appears on the basis of the items used here that students' behavioral values are organized in terms of personal or interpersonal factors, competence or ethical-moral factors, and subject-self or object-self factors. To what extent this would continue to be true with other subject samples and other sets of items remains to be seen. New sets of items are currently being collected in the same open-ended way from different groups of subjects varying in age and occupational status. Similar groups will then be administered the questionnaires to ascertain how pervasive is this organization of behavioral value statements.

#### REFERENCES

1. ARONFREED, J. *Conscience and Conduct: The Socialization of Internalized Control over Behavior*. New York: Academic Press, 1968.
2. BARLOW, J. A., & BURT, C. The identification of factors from different experiments. *Brit. J. Stat. Psychol.*, 1954, 7, 52-56.
3. BASS, B. Social behavior and the orientation inventory: A review. *Psychol. Bull.*, 1967, 68, 250-292.

4. BRIM, O. G., JR. Adolescent personality as self-other systems. *J. Mar. & Fam.*, 1965, **27**, 156-162.
5. CRISSMAN, P. Temporal change and sexual difference in moral judgments. *J. Soc. Psychol.*, 1942, **16**, 29-38.
6. ———. Temporal change and sexual difference in moral judgments. *Univ. Wyoming Publ.*, 1950, **15**, 57-68.
7. FRIEDLANDER, F. Underlying sources of job satisfaction. *J. Appl. Psychol.*, 1963, **47**, 246-250.
8. GOODMAN, N. Adolescent norms and behavior. *Merrill-Palmer Quart. Behav. & Devel.*, 1969, **15**, 200-211.
9. MCKINNEY, J. P. The development of values—Prescriptive or proscriptive? *Hum. Devel.*, 1971, **14**, 71-80.
10. OSGOOD, C. E. Studies on the generality of affective meaning systems. *Amer. Psychol.*, 1963, **17**, 10-28.
11. RESCHER, N. Introduction to Value Theory. Englewood Cliffs, N.J.: Prentice-Hall, 1969.
12. RETTIG, S., & PASAMANICK, B. Changes in moral values among college students: A factorial study. *Amer. Sociolog. Rev.*, 1959, **24**, 856-863.
13. RODGERS, R. R., BRONFENBRENNER, U., & DEVEREUX, E. C., JR. Standards of social behavior among children in four cultures. *Internat. J. Psychol.*, 1968, **3**, 31-41.

*Department of Psychology*

*Michigan State University*

*East Lansing, Michigan 48823*

## THE IMPACT OF EXAM QUESTION ORDER EFFECTS ON STUDENT EVALUATIONS\*

*College of Business Administration, University of Cincinnati*

---

MICHAEL L. DEAN

---

### SUMMARY

After an exam in which the order of question presentation was varied on the basis of pre-established difficulty levels, student perceptions of the test, course, and instructor were compared. Perceptual scores for 68 Ss across nine evaluative factors were analyzed. The hypothesized existence of a primacy/recency halo effect on perceptual bias received partial support. Five of nine factors were significant in the hypothesized direction at the .05 level.

---

### A. INTRODUCTION

The research on serial order effects is not clear-cut; however, there does appear to be general support for the presence of primacy and/or recency effects (3, 4). While there are some dissenting studies in this area (1, 5), the question of which effect is dominant and under what conditions seems to be more the issue than the existence of this learning phenomenon (2).

As an extension of this literature, the subject research was undertaken to determine if order effects, in turn, influence perception. If so, then sequencing may not only influence remembering, but may also affect attitudes. Of particular interest in the research reported here was investigating (a) the opportunity for an instructor to enhance the classroom learning situation and (b) the ability of an instructor through question sequencing to manipulate Ss' perceptions.

### B. METHODOLOGY

#### 1. Subjects

Convenience samples of 81 and 68 undergraduate students, respectively, in introductory business courses at the University of Cincinnati were used. The larger sample provided baseline information for sequencing multiple-

---

\* Received in the Editorial Office on August 2, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.



choice questions according to performance difficulty, while the smaller sample was used as the test group.

## 2. *Materials and Procedure*

Prior classroom examination experience with the same instructor enabled a rank ordering of test questions in terms of difficulty on the basis of the percentage correct response. Two exams were then constructed consisting of the identical 40 questions, but differing in the order of questions. Form A positioned the easier questions in both the first and last quartiles (i. e., 1-10; 31-40), with the more difficult questions appearing in the middle half (i.e., 11-30). Conversely, Form B had the opposite ordering, with the more difficult questions appearing first and last and the easier questions in the middle. In both cases, questions progressed sequentially according to their ranking, either from easiest to most difficult to easiest (Form A) or *vice versa* (Form B). Exams were handed out in a systematic manner, alternating forms. Immediately after turning in his exam, each S was given an evaluation questionnaire which was then submitted anonymously with only Form A or B designated.

It was hypothesized that the existence of a primacy/recency recall effect would influence students' opinions about the exam, the course, and instructor, depending on the particular test version. It was conjectured that the exam form with the easiest questions appearing first and last would produce the most positive perceptual effects. Conversely, it was thought that more negative ratings would result from the forms that positioned the most difficult questions first and last. Likert-type rating scales were used to record opinions toward the fairness and ease of the exam and test score expectation; the course organization, materials, and value; and instructor's ability to stimulate interest, knowledgeability, and overall rating.

## C. RESULTS

Inasmuch as the independent variable was developed in an earlier time frame and represented a potentially contaminating effect, it was important initially to determine the applicability of the pretest questions prior to analyzing comparative test form evaluations. The rank order correlation coefficient between pretest and test ordering of questions was .83, accounting for about 69 percent of the variance and significant at the .001 level. Further, chi square comparisons of Ss' performance for the test forms revealed significant differences at the .05 level between relatively easy (91 percent correct) *versus* difficult questions (68 percent correct). Finally, there were

no differences in the performance scores of the two test groups on the examination itself. Mean differences between Group A and Group B were not significant at the .05 level, suggesting group comparability.

The hypothesis that question ordering would influence Ss' evaluations according to a primacy/recency effect received partial support. As shown in Table 1, of a total of nine evaluative factors, test differences between

TABLE 1  
t TEST COMPARISONS MADE AMONG EVALUATIVE FACTORS FOR DIFFERENT  
QUESTION ORDER TREATMENT

Factor area	Mean rating		Standard deviation	
	Form A	Form B	Form A	Form B
Examination				
Fairness*	1.485	2.000	.225	.287
Ease/difficulty	3.152	3.114	.161	.149
Grade expectation	2.091	2.086	.230	.178
Course				
Organization	1.364	1.543	.191	.177
Materials*	1.667	2.029	.246	.248
Worthwhileness*	1.576	1.771	.250	.155
Instructor				
Superior/inferior*	1.727	1.943	.198	.187
Interest/stimulation	1.636	1.857	.221	.219
Understandable*	1.394	1.829	.208	.338

Note: Ratings range from 1 (most positive) to 5 (most negative).

\*  $p < .05$ .

groups are significant at the .05 level for five factors. Those Ss who had Form A (easy questions clustered at the beginning and end) tended to view the exam as fairer than Form B respondents; however, there were no differences in test score expectations nor perceived levels of exam difficulty between the groups.

Form A Ss felt more positively about the course, tending to view it as more worthwhile and with better course materials than the comparison group. Perceptions of the instructor also tended to be more favorable among Form A respondents. They regarded the teacher as easier to understand and overall more capable.

#### D. DISCUSSION

There is an indication, then, that primacy/recency effects influenced students' perceptions of such classroom factors as the examinations, the

course itself, and the instructor. The implications of these tentative findings are several. It appears that an instructor may be able to enhance the learning experience by judiciously positioning questions so as to cluster the easier ones at the beginning and end with the more difficult ones in the middle. However, inasmuch as this strategy had no differential effect on students' grade expectations, it cannot be concluded that this approach affects an S's confidence in his own performance ability. While enhancing the classroom climate is an assumed desirable effect of clustering, it is possible that such a tactic could artificially inflate student instructor-course evaluations in those cases where evaluations are given immediately following, say, a final exam.

A debriefing of Ss one week after the test provided insights into a possible contamination of effects. It appears that students' strategies in taking multiple-choice exams differ. While most seem to progress sequentially from start to finish, some students answer the (perceived) easiest questions first and then return to the (perceived) more difficult. To the extent that this occurred, then the original ordering of questions was distorted. It should also be noted that this research assumed that easier question recall equates with positive attitudes, whereas recall of difficult questions translates to negative impressions.

While findings in this study do not present an unequivocal picture, they do suggest implications of primacy/recency effects on evaluations. A more conclusive statement awaits further research.

#### REFERENCES

1. BEAN, L. L., & McCROSKERY, J. H. Implications for associative processes of switching the middle of the list during serial rote learning. *J. of Psychol.*, 1973, 83, 227-235.
2. COHEN, A. *Attitude Change and Social Influence*. New York: Basic Books, 1964.
3. EBENHOLTZ, S. M. Serial learning: Position learning and sequential associations. *J. Exper. Psychol.*, 1963, 66, 353-362.
4. MILLER, N., & CAMPBELL, D. P. Recency and primacy in persuasion as a function of the timing of speeches and measurements. *J. Abn. & Soc. Psychol.*, 1959, 59, 1-9.
5. STERNTHAL, B. *Persuasion and the mass communication process*. Unpublished Doctoral dissertation, The Ohio State University, Columbus, 1972. Chap. 8.

*College of Business Administration*  
*University of Cincinnati*  
*Cincinnati, Ohio 45221*

## PERSONALITY AND ATTITUDES: A RE-EMPHASIS UPON THE COGNITIVE COMPONENT\*<sup>1</sup>

*University of Delaware*

---

SALLY M. DUNLAP,<sup>2</sup> SAMUEL L. GAERTNER, AND  
A. DAVID MANGELSDORFF<sup>2</sup>

---

### SUMMARY

Although attempts to demonstrate a relationship between one set of individual differences and the evaluative or behavioral components of an attitude usually have failed, this study examined the possibility that individual differences (i. e., personality) may be important in determining the cognitive component of the attitude (i. e., the rational support offered for the evaluative position or behavior).

Students with idealistic value systems (i.e., justice, egalitarian, humanitarian oriented) endorsed idealistic justifications for their attitudinal position (i.e., pro or anti) toward the war in Vietnam. Pragmatic students (i. e., utilitarian, outcome oriented) endorsed the more pragmatic justifications supporting their attitude toward the war.

---

### A. INTRODUCTION

Most studies that relate personality or attitudinal individual differences to other attitudes have examined the relationship between these individual differences and the evaluative and/or behavioral component of the other attitude. Numerous examples are found in the literature: high authoritarians are more likely to have hawkish attitudes toward the Vietnam war (5), more likely to be antiblack (1), more likely to submit to authority in the Milgram obedience paradigm (4).

Often, however, attempts to demonstrate a relationship between one set of individual differences and the evaluative or behavioral components of an attitude have failed. In his review of the attitude-behavior literature Wicker

---

\* Received in the Editorial Office on August 2, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Requests for reprints and copies of the scales used in the study may be obtained from the senior author at 2300 Nacogdoches Road, 167S, San Antonio, Texas 78209.

<sup>2</sup> Now at Behavioral Science Division, Academy of Health Sciences, Fort Sam Houston, Texas 78234.



(9) reported that the average correlation between attitude and behavior is .30. Wicker's finding is consistent with studies showing strong situational effects on evaluative and behavioral components: e. g., proximity to *E* or the learner determines how far *S* will proceed in the Milgram situation (6). When compared, the situational influence on attitudes/behavior is typically more potent than the individual differences influence (7).

The common interpretation of such comparisons is that individual differences are of minimal importance in accounting for other attitudes and behavior. The authors' contention is that this interpretation may be only partially true because it ignores the possibility that individual differences may be important in determining the third component of the attitude: namely, the cognitive component.

Although situational constraints are potent modifiers of a person's expressed evaluative position and/or behavior, individual differences may still be related to differences in the rational support offered for this position or behavior. For example, a Milgram *S* might rationalize shocking the victim for authoritarian reasons ("the experimenter told me to") or for nonauthoritarian reasons ("I didn't know how else to help him to learn"). While the overt behavior was strongly influenced by the situation, the reasons given for the behavior may reveal the influence of individual differences.

Our approach is theoretically supported by Rosenberg's (8) findings on the consistency between the evaluative and cognitive components of attitudes. Rosenberg's findings imply that cognitively weighed values and perceived instrumentalities lead to evaluative attitudes. Our position maintains the evaluative-cognitive consistency, but shifts the directional emphasis. That is, situationally influenced evaluative attitudes and behavior may lead to varied cognitive supports derived from individually different values and belief systems.

In testing this formulation, the Vietnam war was chosen as the attitude object because of its saliency during the time the study was conducted. Various justifications, for and against the war, were assumed to be familiar to most college students.

The individual differences dimension, idealism-pragmatism (I-P), was chosen because it appeared to be logically related to an apparent split between "moral" and "practical" types of rationales for pro- and antiwar attitudes. Idealism-pragmatism is based on individual differences in values. The idealistic value system is based on ethical principles, such as justice, democracy, and, often, humanitarianism; the pragmatist has a utilitarian value system, emphasizing the ultimate practical outcome (3). As a consequence of the differing values, idealists hold idealistic values as more



important than practical considerations, while pragmatists evaluate actions primarily in light of practical consequences.

Two predictions regarding the manner in which idealism-pragmatism might have been related to Vietnam war attitudes during the Spring of 1971 were derived from the theoretical formulation stated above.

1. *IP individual differences—pro- or antiwar evaluative direction.* IP may have been related to the evaluative direction of the attitude; e. g., idealists were more likely to be doves. Previous evidence suggested that this relationship might exist. Idealists typically report being more left wing politically than pragmatists and more frequently report participation in protest demonstrations than pragmatists. In addition, the stereotypic dove argument that "the war is immoral" is consistent with postulated idealistic tendencies to judge by ethical standards. On these bases the relationship between idealism-pragmatism and evaluative position toward the Vietnam war was predicted. Idealists should have been more dovish than pragmatists.

2. *IP individual differences—justifications used for cognitive support.* IP may have been related to the type of cognitive support offered for the evaluative position. That is, the same evaluative position—e. g., antiwar—might have been justified by idealists as instrumental to idealistically valued ends, but by pragmatists as instrumental to pragmatic ends. Thus idealists might have been antiwar because the war was immoral, or because it interfered with the South Vietnamese right to self-determination. The antiwar position could have been justified by pragmatists on the grounds that it was economically costly, or that it was alienating our former allies.

Similarly, idealists might have been prowar to ensure the South Vietnamese a free choice of government or to protect the South Vietnamese from Communism. In contrast, a prowar attitude might have been justified by pragmatists on the grounds that it prevented the spread of Communism to ourselves, or that it was economically profitable. This second type of relationship was also predicted. Regardless of their evaluative position both idealists and pragmatists would justify their own attitudes toward the war with appropriate value-related reasons. Idealists would use idealistic justifications; pragmatists would use pragmatic justifications.

## B. METHOD

All subjects (General Psychology students at the University of Delaware, Spring, 1971,  $N = 59$  males, 77 females) took a battery of tests in class which included the idealism-pragmatism (IP) personality inventory and the "War in Vietnam Survey."

The IP inventory is a 20 item Likert scale, with a test-retest reliability of

.81. Some items directly support idealistic or pragmatic interests: e. g., "Modern society would benefit much from more concern for the rights and welfare of others" (idealistic item). Others contrast pragmatic and idealistic interests: e. g., "No matter how good your intentions are, what matters is whether you succeed or fail" (pragmatic item). Pragmatic items are reversed in scoring, so a high total score indicates idealism. Ss scoring above the median are idealists; those scoring below the median are pragmatists.

In a college population ( $N = 174$ ), idealism-pragmatism correlated significantly but slightly with authoritarianism ( $r = -.29$ ), Machiavellianism ( $r = -.32$ ), and Factors I (Tough-Tenderminded,  $r = .38$ ),  $M$  (Practical-Imaginative,  $r = .39$ ), and Q1 (Conservative-Radical,  $r = .30$ ) of Cattell's 16 Personality Factor inventory. Pragmatists tended to be authoritarian, Machiavellian, tough-minded, practical, and conservative. These correlations confirm the partial theoretical overlap between idealism-pragmatism and the traits represented by the scales. Failure to obtain correlations above .20 between IP and the remaining 16 PF factors or social desirability (2) which bear no theoretical relationship to idealism-pragmatism, provides evidence of discriminant validity.

In the Vietnam survey Ss first indicated their attitude toward the war as against, supporting, or undecided. Then each  $S$  was asked to rank two lists of possible justifications. Five prowar justifications comprised the first list; eight antiwar justifications comprised the second list. The Ss were to rank the prowar items in the first list as they would if they were truly hawks, in order of their perceived merit as justifications. Similarly, they were instructed to imagine dovish attitudes when they ranked the anti-Vietnam war items in order of their perceived merit. Thus, both hawks and doves ranked possible justifications for their own true attitude and justifications for their counter-attitudinal position. Write-in justifications were permitted, but few students added other reasons.

The two lists of justifications represented commonly heard reasons for holding each attitude during the spring of 1971. Both sets of justifications were ranked in terms of idealism-pragmatism by eight graduate student-faculty judges, using criteria provided by the authors.

The Ss' prowar and antiwar justification rankings were weighted on the basis of the judges' median ranks and summed, so that a high justification score was the result of favoring idealistic justifications. A low justification score resulted from ranking pragmatic justifications higher. Each  $S$  had a prowar justification score and an antiwar justification score.

## C. RESULTS AND DISCUSSION

As predicted, idealists and pragmatists differed in their attitudes toward the war (Table 1). Antiwar Ss (doves) were more likely to be idealistic; undecided Ss and Ss who support the war (hawks) were more likely to be pragmatic ( $\chi^2 = 7.91$ ,  $df = 2$ ,  $p < .02$ ). This result does not necessarily indicate a causative relationship, of course, but the results do demonstrate a relationship between IP individual differences and the evaluative component of Vietnam war attitudes.

TABLE 1  
NUMBER OF IDEALISTS AND PRAGMATISTS WHO WERE HAWKS, DOVES, AND UNDECIDED ABOUT THE WAR, AND CORRELATIONS BETWEEN IDEALISM-PRAGMATISM (IP) AND JUSTIFICATION SCORES

Personality and type of justification	Hawk (support)	Undecided	Dove (antiwar)
<i>Number of idealists and pragmatists</i>			
Idealists <sup>a</sup>	6	13	49
Pragmatists	10	25	33
<i>Correlation between IP and justification scores</i>			
Prowar	.70 <sup>b**</sup>	-.11	.19
Antiwar	.61 *	-.12	.41 <sup>b**</sup>

<sup>a</sup> Idealists scored above the median (100) on IP 5.

<sup>b</sup> Correlation between IP score and justification of S's own attitude position.

\*  $p < .05$ .

\*\*  $p < .01$ .

Prediction 2, that Ss' own attitudes would be supported by appropriate idealistic or pragmatic justifications, was tested by correlating subjects' idealism-pragmatism scores with their justification scores for their own attitudes. A high correlation indicates that idealists favored idealistic justifications and pragmatists favored pragmatic justifications. The appropriate data are the prowar justifications of hawks and the antiwar justifications of doves. As shown on Table 1, for hawks the correlation between IP personality score and war-justification score was .70 ( $df = 14$ ,  $p < .01$ ). For doves the  $r$  was .41 ( $df = 80$ ,  $p < .01$ ). With both hawks and doves, idealists tended to endorse idealistic justifications, while pragmatists endorsed pragmatic justifications of their own attitudes, confirming prediction 2. Idealists and pragmatists differ in the "why" of their attitudes, even when they share the same attitude direction superficially.

This finding, however, is open to the alternative explanation that these correlations are really due to similarities in the wording of the IP scale items and the justifications provided, rather than to a true relationship between

personality and attitudes. To answer this criticism, the IP-justification correlations when Ss were role-playing counter-attitudes were compared to the correlations reported above when Ss were justifying their own attitudes. When doves role-played prowar attitudes, the correlation was .19, compared to .41 when they ranked justifications of their own attitude. When hawks role-played an antiwar attitude, the IP-justification correlation was .61 ( $df = 14$ ,  $p < .05$ ), again tending to be less than when hawks ranked own-attitude justifications ( $r = .70$ ). When Ss were not ranking justifications of their own attitude, trends of lower relationships between IP and justification scores were obtained.

For the Ss who were "undecided about the war" neither set of justifications related to their own attitudes, and hence both justification lists required the adoption of a counter-attitudinal position. The correlations for these "undecided" Ss between IP and pro and anti justification scores, respectively, was  $-.11$  (n.s.) and  $-.12$  (n.s.).

In three of four counter-attitudinal cases (i. e., doves-pro list, hawks-anti list, undecided-pro and anti lists) there was no significant relationship between idealism-pragmatism and the type of justification endorsed. The Ss were consistent in choosing justifications appropriate to their personalities only when justifying their own attitudes. Thus, it seems doubtful that the relationships between IP and own-attitude justifications were merely due to similarity between IP and justification items.

The deviant finding, that hawks showed IP-justification consistency when justifying anti-war attitudes, suggests that hawks were successful at role-playing their counter-attitudinal position. Because of social pressure from doves, hawks may actually have spent more time considering "legitimate" (i. e., cognitive consistent) justifications of the more popular dovish position. It is also possible that Ss who checked the support response were not true hawks but were instead indicating support for President Nixon's graduated withdrawal plan to end the war. For such Ss both pro- and antiwar justifications might have been appropriate to their own true attitudes, and high personality-justification correlations would have been expected for both lists. Unfortunately, no data are available to test either of these explanations of the deviant hawk finding.

The results supporting prediction 2 suggest that the effect of personality on other attitudes may be important even where situational influences are strong. That social climate is important is implied by the fact that over 60% of the Ss reported antiwar attitudes (Table 1). However, within the antiwar group (as well as the prowar group), the IP personality dimension was pre-



dictably related to the reasons endorsed when justifying one's own attitudinal position. Thus, even when situational constraints operate to modify the expression of the evaluative component of attitudes or overt behavior, individual differences may still be evident in the cognitive component of attitudes. These findings suggest a change in strategy for researchers who investigate the relationship between individual differences and other attitudes. That is, it may be more fruitful to focus also on the subtle relationships between individual differences and the cognitive component of the attitude, rather than focusing solely on the more obvious but often limited relationships to the evaluative or behavioral components of the attitude. The results also have implications for attitude change strategies. That is, the results suggest that the content of an attitude change communication should be tailored to the values of the target person. In terms of the present study, arguments pertaining to the immorality of the war would have appealed to idealistic hawks, while those emphasizing the possibility of an improved economy would have been more likely to convert a pragmatic hawk.

#### REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWIK, E., LEVINSON, D. J., & SANFORD, R. N. *The Authoritarian Personality*. New York: Harper, 1950.
2. CROWN, D. P., & MARLOWE, D. *The Approval Motive: Studies in Evaluative Dependence*. New York: Wiley, 1964.
3. DUNLAP, S. M. Reactions of idealists and pragmatists to expectancy disconfirmation: Who will change society? Paper presented at the meeting of the Eastern Psychological Association, New York, 1971.
4. ELMS, A. C., & MILGRAM, S. Personality characteristics associated with obedience and defiance toward authoritative command. *J. Exper. Res. Personal.*, 1966, **1**, 282-289.
5. IZZETT, R. R. Authoritarianism and attitudes toward the Vietnam War as reflected in behavioral and self report measures. *J. Personal. & Soc. Psychol.*, 1971, **17**, 145-148.
6. MILGRAM, S. Some conditions of obedience and disobedience to authority. *Hum. Relat.*, 1965, **18**, 57-76.
7. PETTIGREW, T. F. Regional differences in anti-Negro prejudice. *J. Abn. & Soc. Psychol.*, 1959, **59**, 28-36.
8. ROSENBERG, M. J. Cognitive structure and attitudinal effect. *J. Abn. & Soc. Psychol.*, 1956, **53**, 367-372.
9. WICKER, A. W. Attitudes vs. action: The relationship of verbal and overt behavioral responses to attitude objects. *J. Soc. Issues*, 1969, **25**, 41-78.

*Department of Psychology*  
*University of Delaware*  
*Newark, Delaware 19711*





## COMPONENTS OF DEPRESSION IN ATTEMPTED SUICIDES\*

*Philadelphia General Hospital<sup>1</sup>*

---

AARON T. BECK AND DAVID LESTER

---

### SUMMARY

A factor analysis of the responses of suicide attempters to the Beck Depression Inventory yielded factors similar to those found in previous studies of nonattempters. The consistent association of hopelessness and suicidal wishes in factor analytic studies supports the conclusions from other investigations that hopelessness is an important precursor to suicidal behavior.

---

### A. INTRODUCTION

The term "depression" has many meanings. It may refer to a patient's mood, or it may be a psychiatric diagnosis (3). The derivation of objective rating scales and inventories to assess depression has made it clear that many behaviors enter into an evaluation of depression, however the term may be interpreted. The advent of objective rating scales enables statistical techniques, such as factor-analysis, to be applied to the resulting data in order to ascertain possible objective dimensions of depression.

Many of the factor analytic studies of objective rating scales for depression have used scales that are completed by the psychiatrist or psychologist. Self-report schedules, either administered by an interviewer or self-administered, have been widely used in recent depression research and have been subjected to factor analysis. Analyses of patient-completed scales, such as the Beck Depression Inventory (3), provide the opportunity for comparing the patterns of depressive symptomatology in diverse populations.

Six factor analytic studies of the Beck Depression Inventory have been reported (2, 4, 6, 7, 8, 9). The sample in one of these studies consisted of 606 consecutive admissions to the psychiatric service of a general hospital and to a psychiatric outpatient department (2). The sample included a broad spectrum of neurotic and psychotic disorders, and the inventory was administered by a technician. The samples in the other studies consisted exclusively of depressed patients and the inventory was self-administered.

---

\* Received in the Editorial Office, on August 3, 1973, and published immediately at Provincetown, Massachusetts. Copyright by the Journal Press.

<sup>1</sup> This research was supported by NIMH grant 19989-03.

Despite variations in loadings of specific items, certain regularities occurred among the cited studies. Three factors that consistently emerged were: (a) *Negative Attitudes—Suicide*, defined by pessimism, suicidal wishes, sense of failure, self-accusations, and self-dislike; (b) *Physiological*, defined by anorexia, weight loss, and sleep disturbance; and (c) *Performance Difficulties*, defined by work inhibition and fatigability.

In view of the strength of the factor linking negative attitudes about the self and future with suicidal wishes, it seemed useful to replicate these studies with a sample of attempted suicides. Although the categories of attempted suicide and depression overlap, it is of interest to explore the generality of the factors identified with the use of other criteria for selection of subjects. Furthermore, it is of interest for the study of suicidal behavior to explore whether there are distinct independent dimensions of depression in such individuals.

## B. METHOD

The subjects consisted of the first 254 consecutive attempted suicides admitted to a hospital for this reason. The sample included 114 males and 140 females. All patients were seen within 48 hours of admission on two separate occasions. During the first interview a comprehensive history was taken, and during the second interview a number of psychological tests were administered, including the Beck Depression Inventory.

Data from seven subjects were discarded, since their responses to the Inventory were not complete. The data from the remaining subjects were subjected to a principal components analysis with the use of the factor-analysis program BMD03M of the UCLA Biomedical Statistical Programs package. Five of the eigenvalues were greater than unity, and so it was decided to extract five principal components.

## C. RESULTS

The five factors are shown in Table 1. The first factor has loadings of more than .50 on suicidal wishes, pessimism, lack of satisfaction, crying spells, sad mood, sense of failure, and self-hate. The second factor shows high loadings on items relevant to impaired performance. The third factor, relevant to intrapunitive variables, shows high loadings on sense of punishment, guilty feelings, and self-accusations. The fourth factor is concerned with anorexia and weight loss and the fifth, irritability and somatic preoccupation.

TABLE 1  
THE PRINCIPAL COMPONENTS EXTRACTED FROM THE FACTOR-ANALYSIS AND  
THE LOADINGS FOR THE 21 ITEMS

Item	Component				
	I	II	III	IV	V
Mood	.61	.22	.29	.26	.18
Pessimism	.73	.16	.08	.04	.20
Sense of failure	.57	.18	.27	.14	.20
Lack of satisfaction	.71	.15	.10	.02	.19
Guilty feeling	.64	.09	.51	.02	-.05
Sense of punishment	.23	.22	.71	-.08	-.04
Self-hate	.57	.20	.31	.21	-.15
Self-accusations	.29	-.13	.65	.02	.32
Suicidal wishes	.75	.15	.13	.08	.09
Crying spells	.68	.06	.09	.03	.02
Irritability	.21	.10	-.03	-.04	.69
Social withdrawal	.44	.32	-.12	.22	.21
Indecisiveness	.37	.61	.18	-.03	-.09
Body image	.42	.33	.46	.14	-.01
Work inhibition	.44	.56	.01	-.06	.18
Sleep disturbance	-.06	.12	.48	.20	.46
Fatigability	.19	.67	.17	-.07	.26
Appetite	.12	.30	.14	.71	.20
Weight loss	.15	-.11	-.04	.79	-.04
Somatic preoccupation	.11	.10	.12	.10	.70
Loss of libido	.03	.66	.03	.23	.06
% of total variance	32	7	6	6	5

#### D. DISCUSSION

The findings are similar to previous reports in showing a relationship of suicidal wishes to such cognitive elements as negative expectancies (pessimism) and negative view of the self (sense of failure). There is, however, a stronger representation of affective items in this factor than in the corresponding factor obtained in the studies of patients who had not made recent suicidal attempts. The other factors in the present study roughly parallel those reported in other studies.

The association between negative outlook and suicidal wishes supports recent studies suggesting a causal connection between hopelessness and suicide (1, 5). The present study also suggests that affective symptoms may also contribute to this relationship.

#### REFERENCES

1. BECK, A. T. Thinking and depression: 1. Idiosyncratic content and cognitive distortions. *Arch. Gen. Psychiat.*, 1961, 9, 324-333.
2. BECK, A. T., & OHANESSIAN, A. Assessment of depression: The Depression Inventory. In Pichot, P. (Ed.), *Psychological Measurement*. Basle, Switzerland: Karger, in press.

3. BECK, A. T., WARD, C. H., MENDELSON, M., MOCK, J., & ERBAUGH, J. An inventory for measuring depression. *Arch. Gen. Psychiat.*, 1961, **4**, 561-571.
4. CROPLEY, A. J., & WECKOWICZ, T. E. The dimensionality of clinical depression. *Austral. J. Psychol.*, 1966, **18**, 18-25.
5. MINKOFF, K., BERGMAN, E., BECK, A. T., & BECK, R. Hopelessness, depression, and attempted suicide. *Amer. J. Psychiat.*, 1973, **130**, 455-459.
6. PICHOT, P., & LEMÉRIÈRE, T. Analyse factorielle d'un questionnaire d'auto-évaluation des symptômes dépressifs. *Rev. Psychol. Appl.*, 1964, **14**, 15-29.
7. PICHOT, P., PIRET, J., & CLYDE, D. J. Analyse de la symptomatologie dépressive subjective. *Rev. Psychol. Appl.*, 1966, **16**, 103-115.
8. WECKOWICZ, T. E., MUIR, W., & CROPLEY, A. J. A factor analysis of the Beck Inventory of Depression. *J. Consult. Psychol.*, 1967, **31**, 23-28.
9. WECKOWICZ, T. E., YONGE, K. A., CROPLEY, A. J., & MUIR, W. Objective therapy predictors in depression: A multivariate approach. *Monograph. Suppl.* No. 31, Clinical Psychology Publishing Company, 1971.

429 Stouffer Building

Philadelphia General Hospital

700 Civic Center Boulevard

Philadelphia, Pennsylvania 19104



## CONCEPTUAL SYSTEMS AND PHILOSOPHICAL ORIENTATION\*

*University of North Dakota; and Lakehead University, Canada*

---

JOHN D. WILLIAMS AND JOHN H. KELLEHER

---

### SUMMARY

Subjects ( $N = 152$ ) were administered the Conceptual Systems Test (CST) and the Ross Educational Philosophical Inventory (REPI). The subjects were then classified into one of four conceptual systems, or considered to be an admixture. An analysis of variance was then performed on each of the REPI scales: Idealism, Realism, Pragmatism, and Existentialism. Significance was found among the conceptual systems on each philosophical subtest. System 1 subjects (who are most concrete on a concrete-abstract continuum) had the highest mean on the Idealism subtest; System 2 subjects (who are only slightly more abstract than System 1 subjects) were higher on the remaining three philosophical subtests. Also, the highest correlation occurring among the CST and the REPI occurred between Divine Fate Control and Idealism ( $r = .53$ ).

---

### A. INTRODUCTION

While psychological research has often focused upon various measures of personality, and a similar interest has been focused on philosophical systems by researchers in a philosophical vein, few researchers have attempted to relate the two domains. One of the few attempts was made by Tikalsky (8) who showed that, because of a strong pragmatic vein in American psychology, no strong relationships existed among psychologists regarding preferred philosophical system and area of psychological interest.

Perhaps the most extensive theoretical viewpoint on conceptual systems was made by Harvey *et al.* (2). Most of Harvey's (1) work within the conceptual systems framework has been related to testing the validity of the theory. In two personal applications to the educational setting, Harvey *et al.* (4) found that more abstract teachers differed from the more concrete teachers in directions that are considered desirable. In a followup to that study, Harvey *et al.* (3) found that the students' performance was affected

---

\* Received in the Editorial Office on August 6, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

by the teachers' conceptual systems in ways similar to the first study; i.e., students of abstract teachers were found to be more successful. Other investigators have also used conceptual systems theory in educational applications. Moellenburg and Williams (5) found that superintendents reacted to teacher negotiations in directions predicted by conceptual systems theory. In a study concerning teaching styles, Murphy and Brown (6) found that as the abstractness of the teacher's conceptual system increases, teaching styles are characterized by more sanctioning of search behavior. Less questioning for precise answers takes place.

### B. DESCRIPTION OF THE CONCEPTUAL SYSTEMS

In formulating their theory of conceptual systems, Harvey *et al.* (2) postulated four basic systems, each representing a nodal point along a continuum of the attribute in question. Thus, Conceptual System 1 represents the highest degree of concreteness on the concrete-abstract dimension, followed in order by Systems 2, 3, and 4, with 4 representing the highest degree of abstractness. On other dimensions, the relative order of the systems varies, so that each system represents a constellation of attributes from several dimensions. A brief description of each system may help to clarify this situation for the reader.

Harvey *et al.* (2) describes Conceptual System 1, corresponding to the concrete end of the continuum, as being characterized by a tendency to function in terms of absolutes. Individuals operating in terms of this system can be expected to regard issues categorically as good or bad, true or false, and to base their judgments to a great extent on statements of authority or on rules. A high degree of ethnocentrism and the frequent use of platitudes or conventional statements are also common to this group. Harvey (1) states that their general mode of functioning seems to be closely related to the syndrome of authoritarianism.

Subjects who function predominantly in the pattern of System 2 can be expected to exhibit a high degree of negativism and to denounce institutional authority. In this respect, they are directionally opposite from the System 1 individuals who rely on authority. However, their conceptual systems are quite similar in structure, in that System 2 is also characterized by a high degree of dogmatism or absolutism. These individuals are quite determined in their drive toward autonomy, and emphatically reject any dependence upon authority.

System 3 is the classification used for subjects who place a great deal of stress on interpersonal relationships and acceptance by their peers. These

individuals are less absolutistic than either of the previously described types in their mode of thought, since System 3 falls next to the most abstract position on the concrete-abstract continuum. In their concern for peer acceptance, they are prone to conform to the attitudes and values of significant others, and to avoid issues that might cause rifts and resulting loneliness or isolation.

System 4, the highest level of abstractness posited by Harvey *et al.* (2), is the category for individuals who are able to maintain a high level of independence without the negativism found in System 2, and are able to base their decisions upon informational standards rather than upon authoritarian pronouncements or dependency relationships. Their responses to problem situations indicate a high degree of novelty and appropriateness, with a general usage of multidimensional rather than unidimensional criteria (1).

### C. PROCEDURE

Subjects were 152 undergraduate students preparing to be elementary teachers at the University of North Dakota. Each subject was administered the Conceptual Systems Test (CST) and the Ross Educational Philosophical Inventory. The CST was developed by Harvey and others (1) and was devised to be consonant with conceptual systems theory (2). The CST has six subtests: Divine Fate Control, Need for Structure-Order, Need for People, Need to Help People, Interpersonal Aggression, and Anomie.

Ross (7) developed his inventory to determine a person's philosophical orientation and the strength with which it is held. Items on the REPI are ontological, epistemological, and axiological statements from the following philosophical systems: Idealism, Realism, Pragmatism, and Existentialism. These four philosophical systems are also the four subscales of the REPI. Both the CST and the REPI use a Likert scale.

### D. RESULTS

Classification into the four conceptual systems was based on the following criteria: using the factor scores on the six scales, a subject was classified a System 1 if the score on Divine Fate Control was 4.19 or higher; System 2 subjects were those whose score on Divine Fate Control was less than 4.19, above 3.75 on Interpersonal Aggression, and above 3.39 on Anomie; to be classified in System 3, the score for Divine Fate Control was below 4.19, below 3.75 on Interpersonal Aggression, and above 4.10 on Need for People; to be classified in System 4, the scores were below 4.19 on Divine Fate Con-

TABLE 1  
ANALYSES OF VARIANCE FOR IDEALISM, REALISM, PRAGMATISM, AND EXISTENTIALISM:  
CLASSIFICATION ACCORDING TO CONCEPTUAL ORIENTATION

Variable and group	N	Mean	Source	df	SS	MS	F	p
<b>Idealism</b>								
System 1	77	63.87	Among	4	1801.26	450.32	11.18	<.001
System 2	4	61.25	Within	147	5918.82	40.26		
System 3	52	58.35	Total	151	7720.08			
System 4	14	53.93						
Admixture	5	56.67						
<b>Realism</b>								
System 1	77	59.43	Among	4	744.93	186.23	3.50	<.01
System 2	4	62.75	Within	147	7824.78	52.23		
System 3	52	57.50	Total	151	8569.71			
System 4	14	53.93						
Admixture	5	51.50						
<b>Pragmatism</b>								
System 1	77	67.08	Among	4	465.38	116.34	2.89	<.05
System 2	4	76.00	Within	147	5925.25	40.31		
System 3	52	68.02	Total	151	6390.63			
System 4	14	65.00						
Admixture	5	66.17						
<b>Existentialism</b>								
System 1	77	69.23	Among	4	700.94	175.24	3.44	<.05
System 2	4	77.25	Within	147	7481.45	50.89		
System 3	52	68.83	Total	151	8182.39			
System 4	14	64.29						
Admixture	5	64.33						

trol, below 4.19 on Need for Structure-Order, below 4.10 on Need for People, and below 3.75 on Interpersonal Aggression; those subjects who did not fit any of the previous criteria were classified as admixtures. Table 1 contains the analysis of variance for each of the four philosophical scales. The philosophical scale of Idealism shows System 1 subjects scoring the highest of the five groups ( $p < .001$ ). For the remaining three scales, the System 2 subjects scored highest. Table 2 contains the correlations between the CST and the REPI. The Idealism subscale is significantly ( $p < .01$ ) related to four of the six CST measures and reflects a similarity to Harvey's lowest level of conceptual organization. The Pragmatism scale shows a general lack of relationship with the CST factors, with the exception of the Anomie scale. On the other hand, the Anomie scale correlated almost uniformly to the same degree with each REPI measure.

### E. DISCUSSION

Prior to the collection of the data, it seemed reasonable to conjecture higher scores for System 1 individuals on Idealism; System 3 individuals on Pragmatism; and System 4 individuals on Existentialism. The reported findings for System 1 subjects were in the hypothesized direction. The generally lower scores of Systems 3 and 4 subjects on the major philosophical systems, while unexpected, were not contradictory to the Conceptual Systems Theory. The lower scores seem to have resulted from abstract subjects making finer discriminations on each of the statements contained in the various REPI scales. As the scores of Admixtures were consistently intermediate to those of Systems 3 and 4 subjects, this suggested a higher level of conceptual development for a considerable proportion of this group.

Of the REPI scales, Pragmatism and Existentialism showed the least relationship to the CST scales. On the other hand, the Idealism scale showed a

TABLE 2  
INTERCORRELATIONS BETWEEN THE CONCEPTUAL SYSTEMS TEST AND  
THE ROSS EDUCATIONAL PHILOSOPHY INVENTORY

Conceptual Systems Test factors	Ross Educational Philosophy Inventory scales			
	Idealism	Realism	Pragmatism	Existentialism
Divine Fate Control	.53*	.25*	.09	.12
Need for Structure-Order	.38*	.30	.16	.02
Need for People	.35*	.11	.15	.21*
Need to Help People	.14	.17	.06	.08
Interpersonal Aggression	-.04	-.01	.12	.10
Anomie	.26*	.25*	.26*	.24*

\* Significant at .01 level.



high positive relationship to four of the six scales; the correlation with Divine Fate Control was the highest ( $r = .53$ ). Evidently, the Idealism scale fit rather well in describing subjects who function in System 1. The other three scales did not show such a correspondence to conceptual systems theory.

#### REFERENCES

1. HARVEY, O. J., *Ed.* Experience, Structure and Adaptability. New York: Springer, 1966.
2. HARVEY, O. J., HUNT, D. E., & SCHROEDER, H. M. Conceptual Systems and Personality Organization. New York: Wiley, 1961.
3. HARVEY, O. J., PRATHER, M., WHITE, B. J., & HOFFMEISTER, J. K. Teacher's beliefs, classroom atmosphere and student behavior. *Amer. Educ. Res. J.*, 1968, 5(2), 151-166.
4. HARVEY, O. J., WHITE, B. J., PRATHER, M., ALTER, R. D., & HOFFMEISTER, J. K. Teachers' belief systems and pre-school atmospheres. *J. Educ. Psychol.*, 1966, 57(6), 373-381.
5. MOELLENBERG, W. P., & WILLIAMS, J. D. Conceptual systems theory and the superintendent in teacher negotiations. *ISR Journal*, 1969, 1(1), 64-78.
6. MURPHY, P. D., & BROWN, M. M. Conceptual systems and teaching styles. *Amer. Educ. Res. J.*, 1970, 7(4), 529-540.
7. ROSS, C. An educational philosophical inventory: An instrument for measuring change and determining philosophical perspective. *J. Educ. Thought*, 1970, 4, 20-26.
8. TIKALSKY, F. D. Values and psychological theory preference. *J. Res. Serv.*, 1966, 6, 3-27.

*The Center for Teaching and Learning*  
*The University of North Dakota*  
*Grand Forks, North Dakota 58201*

## EPISODIC ANALYSIS OF NOVELS\*

*Stanford University*

---

ROBERT R. SEARS AND DEBORAH LAPIDUS

---

### SUMMARY

Episodic analysis is the separation of a novel into its component episodes—i.e., into short stories each of which has a unitary theme—in order to provide equivalent units for content analysis. Independent coders show two types of disagreement. Neither is serious for content analysis, and both may be minimized by some suggested ground rules. In a study of 13 stylistically different books, intercoder reliability ranged from 76% to 94%. Variations in accuracy were clearly related to literary style. Relevant information on the coding process and reliability, as well as data on the number and length of episodes in the 13 books, are given in tabular form. Problems connected with reliable coding are discussed and some solutions are suggested.

---

### A. INTRODUCTION

In connection with a case study of Mark Twain's personality development, a method was needed for assessing the changes in strength of certain of his feelings and motives as he grew from youth to later maturity. The biographical material describing his overt social behavior, and the circumstances under which it occurred, provides a set of data from which one type of assessment can be made. Personal documents, such as letters and autobiographical writings, provide another. The obvious constraints on publicly observable behavior limit direct expression of certain kinds of feelings, however, and self-protective mechanisms against the arousal of anxiety serve to disguise and inhibit others. The data from these sources are obviously essential for a case study, but they cannot provide a complete description of motives and feelings, especially of the defenses which accompany some of them.

In a living person, fantasies—dreams, daydreams, and responses to projective tests—are often used to supplement the more open forms of behavior and self-examination as vehicles for the expression of feelings. For a dead novelist, fantasies come most clearly from his fictional writings. While they

---

\* Received in the Editorial Office on August 6, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

are subject to many controls of a literary nature, unlike the daydreams or test responses of a nonpublishing research subject, they are subject to fewer and to quite different expressive constraints than day-to-day behavior or personal documents. The free flow of narrative is difficult to analyze quantitatively, but for a longitudinal study comparative statements are needed about motive or feeling strengths at different times in the life cycle. An author's novels (or stories, or plays) are successive fantasy-behavior outputs; their motivational contents are subject to the changing experiences of his life and of his own personality development. If the fantasies can undergo quantitative measurement, successive pieces of fiction can be used for determining changes related to these antecedents.

A suitable model for such measurement is the type of content analysis developed by McClelland and his colleagues (13) for assessing achievement motivation from stories told in response to Thematic Apperception Test pictures. The two essential qualities of these stories are that they are (a) brief, and (b) thematically unitary: i.e., centered around a single perceptual event. Their brevity makes them easily subject to content analysis, and their unity of theme makes them equivalent to one another for purposes of counting frequencies and securing average scores regardless of the number of stories used. Novels vary in length, and so do the story themes they encompass. To convert the free-flowing continuity of narrative to a succession of thematic units suitable for content analysis, we have devised an analytic method that breaks the text into episodes that are psychologically equivalent to one another.

### B. DEFINING AN EPISODE

An episode is a small story with a unitary theme, a theme consisting of what the protagonists are doing, saying, thinking, or feeling. The operative word is *what*: i.e., the content of the protagonists' actions. There is always a manifest substantive quality to these verbs; an actor does something, says something, thinks about something, or feels in a certain way, usually about something. The unity consists in the singleness of this *something*. For example, when Tom Sawyer met the new boy in the street, felt his own status threatened, and then had a fight which ended in victory, Mark Twain was clearly defining a unitary event. It started with a new person's being introduced, continued with a certain kind of action (fighting), and ended with the boy's being taken off stage and Tom's being caught and punished by Aunt Polly. There was a definable beginning, which differed substantially from what had

been occurring before (Tom's practicing his new whistle), and an end, which differed from what came next (facing the whitewashing job next morning).

Separating a novel into episodes involves the identification of the beginning of a unitary theme and its end. Although an episode is something constant and continuous and has unity, the actual task of discrimination is one less of constancy than of change. The analyst must identify the point in the exposition at which a new theme is introduced, rather than note continuously the persistence of the current one.

There do not seem to be any universal easily-noted indices that can improve the objectivity of the judgement as to when an episode begins and ends. There are a few rather common correlates of change of theme but none that can be used automatically. A new episode often has new major protagonists introduced or previous ones withdrawn. Often, too, the time or place of the action changes. On occasion, when neither of these occurs, there may be a change of mood, as is exhibited in the cave scene in *Tom Sawyer* (5) at the point when the children suddenly realize they are lost.

More reliable sign-posts to change are those aspects of an action sequence that are commonly used to describe the basic parameters of behavior. *Instigation* is an important determinant of the thematic character of any action sequence. Some of this is external, the actions of other people, physical changes in the environment, the introduction of a new social situation. When Tom and Huck and Joe on Jackson's Island hear the booming of the ferryboat cannon, which is trying to bring their supposedly drowned bodies to the surface, a new episode obviously starts. Or new instigation may be internal, as when Tom's brooding about the loss of Becky's love sends him off miserably to Cardiff Hill.

*Purpose* is another element in much behavior and a change in the protagonist's purpose may signal a new episode. A theme quite commonly involves a series of actions that are directed toward some goal. The goal may be an actor's, as in the case of Tom's perfecting his new whistling technique, or it may be the author's, as when the Connecticut Yankee describes the cruelty of Morgan Le Fay's treatment of the young poacher, a transparent device for arousing the same moral indignation in the reader that Mark Twain himself felt toward royalty, or more broadly, authority.

*Instrumental activity*—i.e., what the actors are doing—described in common-sense terms, almost always changes with the start of a new episode. This would be a useful index if the character of the on-going activity did not change so frequently within episodes also! The goal or purpose binds instru-



mental behavior into a sequence, and the purpose is more likely to be the element which provides thematic unity and continuity to a series of actions than the nature of the actions themselves. In the delicious—and long—episode in *Huck Finn* (7), in which Huck disguised as a girl, visits a strange woman to pick up town gossip, there are a half dozen changes in what Huck does and what the two talk about, yet this is clearly a single episode by the criterion of thematic unity.

The nature of *aroused needs* in the actors, their *emotions*, the *manipulanda* they use in carrying out the action of the story, all change quite commonly when there is a change of theme. If one were to make a list of correlates, from most trustworthy to least, it would include the following changes: (a) purpose or goals of the main protagonists; (b) nature of the activity they perform; (c) external and internal instigation, including thoughts and expectancies; (d) who the main protagonists are; and (e) setting. But it is the theme, not these correlated factors, that finally determines what constitutes an episode.

### C. CODING RELIABILITY

The definition of an episode in terms of thematic unity is notably abstract. In the absence of automatic indices of continuity or change, one reasonable test of validity is the degree of intercoder reliability. To test the effectiveness of the method, as well as the objectivity of our analytic judgements, we coded 13 books by this procedure, working independently, and then counted the number of agreements between our choices of lines for the beginning and end of each episode. Seven of the books are novels by Mark Twain. One of his travel books, *The Innocents Abroad* (4), was added to determine whether the method would be applicable to such writing. The remaining five novels, by Hemingway, Hesse, Steinbeck, and Stevenson (10, 11, 12, 14, 15), were chosen to provide some variety in authorial style. The 13 books are listed in Table 1, which provides data concerning the reliability of our codings by reference to intercoder agreement.

In column A under Reliability are shown the percents of episodes for which exact agreement occurred; i.e., both coders chose the same word with which to mark the end of an episode. The figure was obtained by the formula  $(2 \times \text{no. of agreements}) / (\text{no. of episodes judged by L} + \text{no. judged by S})$ . In the columns labeled L and S are shown the number of episodes into which each of us separated a novel initially. In the column labeled No. of Episodes



TABLE 1  
NUMBER OF CODING JUDGEMENTS, AGREEMENTS, AND ERRORS WITH THREE TYPES OF RELIABILITY MEASURES  
(PERCENT AGREEMENTS), AND NUMBER AND LENGTH OF EPISODES

Book	No. of judgements*		No. of agree-ments	No. of two error types		Reliabilities (% agreement)*			No. of episodes	No. of words per episode	
	L	S		Trans.	Seg.	A	B	C		Range	Mean
Mark Twain											SD
Stormfield (6)	19	13	12	1	6	75	81	100	16	385-1780	938
Tom Sawyer (5)	114	101	77	8	19	72	79	90	81	76-2335	838
Huck Finn (7)	97	94	67	16	25	70	87	100	102	150-3410	1099
Conn. Yankee (1)	100	82	70	10	19	77	88	98	86	215-3690	1356
Pudd'nhead (3)	48	49	36	8	9	75	92	100	49	307-3869	1063
Myst. Stranger (8)	36	30	20	6	11	61	79	95	34	300-2055	1075
Innocents (4)**	96	104	91	3	18	91	94	100	317	38-2554	557
Prince/Fauper (2)	82	65	50	14	19	68	87	100	70	209-2878	980
Hemingway											
Sun Also Rises (11)	66	88	47	13	33	61	78	100	90	152-2200	619
Old Man (10)	48	53	30	13	15	59	85	100	51	175-1245	511
Hesse											
Siddhartha (12)	58	44	29	10	18	57	76	94	49	136-2493	741
Steinbeck											
Mice and Men (14)	43	35	28	4	12	72	82	94	38	203-1757	842
Stevenson											
Treasure Island (15)***	48	54	36	8	14	73	89	100	—	—	—

\* See text for the three formulae for determining reliability. The judges were the two authors.

\*\* Reliability values are based on Chapters 1-20; Chapters 21-40 were coded singly by L, and 41-60 by S.

\*\*\* Two analyses by S, five months apart.

is given our final joint judgment as to the proper division of the text after a line by line examination of our disagreements.<sup>1</sup>

Except for *The Innocents Abroad*, the most stringent test of reliability (A) is not satisfactory for any of the books. The lack of perfect agreement is less serious than these figures might suggest, however. There are three sources of disagreement, and two of these, which constitute the great majority, can be minimized by adoption of a few ground rules.

1. *Transitional errors* are those resulting from a disagreement over the exact sentence at which one episode ends and the next begins. In the Mark Twain novels, these disagreements often involved only one or two sentences and never exceeded a paragraph. Since a transition usually introduces the new thematic topic while closing out the previous one, the sentences contain information about both. A re-examination of the disagreements in these books shows that most of them could have been avoided by following the rule: always attach the transitional segment to the next episode unless this leaves the previous one with no ending whatever.

If transitional errors are disregarded, the intercoder reliability is significantly increased. In column B of Table 1 are shown the percents of agreement for the 13 books by the formula:  $[2 \times (\text{no. of agreements} + \text{transitional errors})] / (\text{no. of judgements by L} + \text{no. by S})$ . Except for *Siddhartha*, which we chose deliberately because Hesse was not an episodic writer in the same sense that Mark Twain was, the B reliabilities are reasonably satisfactory. While one may have greater confidence in the final "discussed" codings than in either of the single coders' judgements, the B reliabilities are high enough to warrant usage of either coder's episodes alone as a basis for a content analysis.

2. *Segmental errors* are disagreements as to how many episodes a given passage in the novel is to be divided into. One coder may judge a certain sequence of 1000 words to represent a single episode, while another coder may break it into two. Obviously this is a more serious kind of disagreement, for it affects the size of both the denominator and the numerator if one is counting relative frequency of some content category. In a book containing 50 or more episodes, these differences produce rather small variations in

<sup>1</sup> A listing of the episodes for each of the books mentioned in Table 1 is available from the authors. The descriptive listing for each episode is by reference to chapter, page number, and opening text phrase in the easily available paperback editions given in the reference list. Paine's highly edited edition of *The Mysterious Stranger* (8) was used for this exercise because of its easy availability. The full text of Mark Twain's own unedited manuscript, *The Chronicle of Little Satan*, (9) has been used for other analyses, and a listing for it is available also.

obtained values from a content analysis, however, and the ground rules suggested below help to reduce the disagreements.

In Table 1, under column C, are shown the intercoder percent agreement values when segmental errors are treated arithmetically as if they were non-errors. They are errors, of course, and the following formula for Column C is used only to discover how much "true error" there is in the coding of a given book:  $[2 \times (\text{no. of agreements} + \text{transitional errors}) + \text{no. of segmental errors}] / (\text{no. of judgements by L} + \text{no. by S})$ .

3. *True errors* are disagreements which stem from the inability of two coders to agree as to what is or is not a unitary theme. The most notable example of this difficulty in Mark Twain's writing is provided by Chapters 32 through 34 of *Tom Sawyer*, the sequence involving the following events: Tom's return from the cave, his talk with Huck, his interactions with Judge Thatcher, the realization (and discovery) of Injun Joe's fate, the interaction with the Welshman, and the beginning of the party at the Widow's. The sequence of events is clear, but one theme tumbles over another so fast that a coder cannot be orderly in his analysis. Similar difficulties occur in the Hemingway and Hesse novels, and even in Steinbeck. The opening two episodes in *Of Mice and Men* are clear enough, for example, but the sensitive intertwining of the mouse, beans and ketchup, and rabbit themes in the ensuing eight pages requires some arbitrary decisions based on the same principles as those used for eliminating segmental disagreements.

#### D. SPECIAL PROBLEMS AND GROUND RULES

##### 1. *Segmental Structure*

An understanding of what constitutes a *theme* has not progressed much since Aristotle considered the issues in *The Poetics*. Our effort to use thematic unity as a criterion for separating episodes has revealed a number of common correlates but no unequivocal indices of the beginning and end of themes when they are embedded in the larger frame of a novel. Hence we have considered the problem of segmental structure from both a theoretical psychological and a pragmatic measurement standpoint.

How long should an episode be? Answer: as long as possible and as short as possible. The *long* answer is the theoretical one. It is predicated on the wish to include in an episode the full expression of a theme. Each episode is intended to be the equivalent to every other. If it is, there is justification for counting the frequency of episodes containing whatever content categories we are studying. We do not want to weight one theme more heavily than

another, as would be the case were one unitary theme segmented into two or more episodes while another was left singular.

The *short* answer is pragmatic. Content analysis—that is, the coding of a textual unit (in this case, an episode) according to presence or absence of predetermined categories—becomes more and more difficult the longer the textual unit is. Content analysis requires a firm focus of attention on every word and line. Attention flags. Words and phrases slip by unheeded, especially if the passage has been read and coded many times before. In our analyses of aggression, exhibitionism, and attachment behaviors in the Mark Twain novels, to be reported later, we have found many slippages, or missed instances, which clearly resulted from this factor, and they were particularly prominent in the longer episodes.

For pragmatic reasons, therefore, we have followed the ground rule: with a seemingly short episode, do not segment it more than is essential; with a seemingly long one, break it if it does not do severe violence to the requirement of thematic unity.

## 2. Chapters

In most cases, the authors of these books have divided their narratives into chapters of which the beginnings and endings correspond to our judgments of the break-points for episodes. There are more episodes than chapters in every instance, however, and one cannot wait for a chapter end to end an episode. Furthermore, while a chapter end may correspond to an episode end, it is not an infallible guide. In *Siddhartha*, for example, the story is separated into a dozen segments by the author, but only two of these separations correspond to our episode separations. At the other extreme, each chapter ending in *Of Mice and Men* terminates an episode. In the seven Mark Twain novels, there are 190 chapters altogether, and 83% of their endings serve to terminate episodes, but in the other 17%, an episode overlaps the break.<sup>2</sup> A suggested guideline is as follows: look sharply at any formal break in the narrative that the author has introduced, but do not rely on it as a hard and fast criterion of a break between episodes.

## 3. Historical Settings and Character Descriptions

Three of these seven novels open with substantial historical reports. *Huck Finn* summarizes the final events related in *Tom Sawyer*, while *Pudd'nhead*

<sup>2</sup> The photocopy of the *Tom Sawyer* manuscript in the Mark Twain Papers shows that the author himself made the chapter separations. Only 6% of these chapter endings fail to correspond with our episode endings. Prof. Henry Nash Smith (personal communication) has pointed out that the *Huck Finn* photocopy is chapterless, however, Mark Twain possibly having left the separation to a copy-editor. This may explain why 25% of the chapter endings in *Huck* do not correspond to episode endings.



*Wilson* (3) and *The Mysterious Stranger* give an account of the settings—historical, geographical, cultural—and the principal characters which will become the scene and the human building blocks for the novels. In each of the three, this introductory material comprises two episodes, the first more arm's length and objective, the second subjective about the narrator or, in the case of *Pudd'nhead*, the titular character. Similar historical vignettes, with character sketches, occur commonly in Dumas, the most notable in *The Three Musketeers* being the fully autonomous histories of Athos, Porthos, and Aramis. In all these instances there is at least superficial confusion caused by the references to different characters and their backgrounds, and to different locales. We have concluded, however, that the history, or the setting, or the character—whichever seems the most central to moving the story forward—should be considered the determinant of the theme. As long as Huck was concerned with the history of his and Tom's adventures in the previous book, or as long as Theodor was describing Eseldorf and the main adult protagonists, or Dumas was describing Athos as a person defined by his experiences and behavior (regardless of the years and geography covered), a unitary theme was being presented.

A special difficulty may arise when there is a flashback in what otherwise seems to be a simple episode, as when the Yankee is reminded, by Sandy's sing-song yarnning about knightly encounters, of the old New Haven Railroad conductor's sing-song announcement of stations. We simply made a ground rule: if the history or other diversion is brief—a hundred words or so—and clearly a subordinate part of a theme that then continues to a normal ending, it is not to be isolated. If it is the length of an average episode itself, however, and contains a clear and independent theme, it becomes an episode.

#### 4. Editorials

Mark Twain found the first-person narrative style most congenial to his purposes. Only three of his major works were written in the third person (*Tom Sawyer*, *Prince and the Pauper* (2), *Pudd'nhead Wilson*). The rest, together with all his travel books and the majority of his short stories, were in the first-person, though in many of these latter the first-person viewpoint was more a story frame than a reflection of the pure introspective style so consistently maintained in *Huckleberry Finn* and *A Connecticut Yankee* (1). These style differences make little difference to episodic analysis except in one respect—the first-person narrative lends itself better to editorializing, and editorials sometimes cause trouble.

A simple instance in the third person is at the end of the whitewashing episode in *Tom Sawyer* (Chapter 2). The author refers to himself directly, and



gives his opinion about what constitutes "work." The editorial is brief, strictly cogent to what Tom has just achieved, and there is no question that it provides a terminating paragraph for a single episode. When we get to the prolonged invectives of *The Mysterious Stranger*, now presented by Satan as an editorial writer for the author, the decision as to what constitutes an episode is less clear. The frequency, length, and bitterness of Mark Twain's editorials increased enormously from *Stormfield* (6) to *The Mysterious Stranger*, but we have thought it better for these qualities to be revealed through content analysis of episodes than through the choice of passages selected to define the episodes themselves. Hence our ground rule is that editorial passages, no matter how long or how infuriated, be included as parts of those episodes for which the manifest story content defines the unity of theme.

## REFERENCES

1. CLEMENS, S. L. (MARK TWAIN). *A Connecticut Yankee in King Arthur's Court* (1889). New York: New American Library, 1963.
2. ———. *The Prince and the Pauper* (1881). New York: New American Library, 1964.
3. ———. *Pudd'nhead Wilson* (1894). New York: New American Library, 1964.
4. ———. *The Innocents Abroad* (1869). New York: New American Library, 1966.
5. ———. *The Adventures of Tom Sawyer* (1876). New York: Bantam Books, 1966.
6. ———. Extract from Captain Stormfield's Visit to Heaven (1907; written in 1868). In C. Neider (Ed.), *The Complete Short Stories of Mark Twain*. New York: Bantam Books, 1966.
7. ———. *Adventures of Huckleberry Finn* (1884). In B. DeVoto (Ed.), *The Portable Mark Twain*. New York: Viking Press, 1968.
8. ———. *The Mysterious Stranger* (1916; written 1897-1904). In B. DeVoto (Ed.), *The Portable Mark Twain*. New York: Viking Press, 1968.
9. ———. *The Chronicle of Young Satan* (1897-1900). In W. M. Gibson (Ed.), *Mark Twain's Mysterious Stranger Manuscripts*. Berkeley: Univ. California Press, 1969.
10. HEMINGWAY, E. *The Old Man and the Sea* (1952). New York: Scribners, 1952.
11. ———. *The Sun Also Rises* (1926). New York: Scribners, 1970.
12. HESSE, H. *Siddhartha* (1922). New York: Bantam Books, 1971.
13. MCCLELLAND, D., et al. *The Achievement Motive*. New York: Century-Appleton-Croft, 1953.
14. STEINBECK, J. *Of Mice and Men* (1937). New York: Bantam Books, 1971.
15. STEVENSON, R. L. *Treasure Island* (1883). New York: New American Library, 1965.

Department of Psychology  
Stanford University  
Stanford, California 94305

## LEXICAL MARKING EFFECTS IN THE SEMANTIC DIFFERENTIAL\* <sup>1</sup>

*Michigan State University*

---

ANDREW R. GILPIN

---

### SUMMARY

The research reported below is directed at the phenomenon of asymmetry in semantic differential scales. It is hypothesized that one source of such asymmetry is the "Lexical Marking Effect" (3, 4). Two experiments are reported which tested this hypothesis. In Experiment I, college students rated concepts on bipolar scales (e.g., Good-Bad) and on unipolar scales involving lexically unmarked (e.g., Good-Not Good) and lexically marked (e.g., Not Bad-Bad) adjectives. Unmarked unipolar ratings resembled bipolar ratings more closely than did marked unipolar ratings, evidence supporting the hypothesis. In Experiment II, similar results were obtained, but the effect was attenuated by requiring subjects to rate concepts rapidly. This suggests that the effects are apparently not due to errors in processing marked terms.

---

### A. INTRODUCTION

Psychologists have long been interested in the extent to which ostensibly antonymous adjective pairs (e.g., good-bad, long-short) are true psychological opposites. Recent interest in this question has focused on the lexical marking phenomenon (8). For many (if not all) such pairs, one member, the marked term (e.g., bad, short) is structurally more complex than the other, unmarked term (e.g., good, long). Clark (3, 4) argued that in order to process information expressed with the use of marked terms, one must transform the information into the equivalent unmarked form. Lexical marking effects were initially demonstrated in solution latency and errors in solving three-term series problems, a task for which this interpretation is not unchallenged (12,

---

\* Received in the Editorial Office on August 7, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The author acknowledges his debt to the late Dr. William T. Stellwagen, who collaborated on some of the pilot work. Earlier versions of this report were criticized by Hiram E. Fitzgerald and Bradley B. Glanville; however, any errors are the responsibility of the author.

13). However, analogous effects have been demonstrated on other tasks: for example, word association (5); numerical answers to questions (10); description of spatial relationships (6); and release from proactive inhibition (20). The lexical marking research provides impressive evidence that antonymous adjectives are not used symmetrically on a number of tasks.

There is at least one context, however, in which antonyms are typically assumed to be symmetric: *viz.*, semantic differential research. Use of the semantic differential technique requires the assumption that the scales whose end points are defined by pairs of antonymous adjectives are in fact symmetric: i.e., the scales extend in opposite directions from neutrality and the end points are equidistant from the neutral point (17). A number of investigators have tested this assumption; while such variables as abstractness of the concepts being rated and subject responses style may affect the degree of symmetry of the scales, results of these studies taken as a group are aptly summarized by Bentler's title (2), "Semantic space is (approximately) bipolar" (1, 7, 11, 14, 15, 16, 18). Although there is a great deal of evidence that many of the scales are not perfectly symmetric, the departures from symmetry have been attributed to extrinsic factors, such as subject variables (2), characteristics of the concepts being rated (1), or to capricious characteristics of some of the scales (7). An alternative explanation would be to attribute asymmetry to structural configuration of the terms defining the scales; and one obvious possibility would be to examine the relationship between asymmetry and lexical marking. (It is important to note that this assertion does not imply that other sources of asymmetry do not exist as well.)

The two experiments reported in the present paper were directed at this issue. The first experiment was an attempt to show that asymmetry is systematically related to lexical marking; the second experiment was an attempt to demonstrate that the asymmetry arises from structural characteristics of the terms rather than from errors in cognitive processing occurring during the rating process *per se*.

Both experiments employed a modification of the Taylor and Kumata (18) unipolar rating scale technique, in which ratings of concepts on bipolar scales are compared with ratings of the same concepts on unipolar scales. It was assumed that ratings of a concept on a conventional bipolar semantic differential scale (e.g., GOOD:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:BAD) would resemble ratings of the same concept on a unipolar scale involving only the unmarked adjective (e.g., GOOD:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:NOT GOOD) more than the former would resemble ratings on a scale involving only the corresponding

marked adjective (e.g., NOT BAD:—:—:—:—:—:—:BAD). This prediction follows from the assumption that scales expressed in terms of the unmarked words can represent the entire bipolar dimension, whereas scales expressed in terms of marked words can represent only half of the bipolar dimension (*cf.* 3, 4, 12, 13).

## B. EXPERIMENT I

### 1. Method

*a. Subjects.* Subjects were 35 students enrolled in an introductory level psychology course at Michigan State University. Participation in the experiment was not a course requirement. All subjects were run in one session. Data from three subjects were discarded: two failed to complete all ratings, and one consistently marked all ratings on one side of the page.

*b. Materials and Procedure.* Each subject rated each of four concepts (candy, animal, psychology, peace) on each of four test scales (good-bad; long-short; clean-dirty; beautiful-ugly); however, each test scale appeared in three forms (unmarked unipolar, marked unipolar, bipolar). In addition, each concept was rated on 12 filler scales, consisting of unipolar and bipolar scales constructed from other semantic differential scales, chosen randomly for each concept. Thus, each subject made a total of 96 ratings. All scales were six-step scales. Booklets consisted of 12 pages, plus an initial instruction sheet; at the top of each page was a concept to be rated (each concept appeared on three different pages), and below it were listed eight scales: four filler scales, and four test scales, arranged in random order. The side of the page on which particular ends of test scales appeared was counter-balanced across concepts. Unipolar and bipolar scales were mixed on each page, but only one of the three forms of a test scale appeared on any one page; pages were assembled in random order. The subjects were given standard semantic differential instructions (17, pp. 82-84) modified for use with six-step scales. All subjects were given as much time as they wished to complete ratings.

### 2. Results

Responses were scored in such a way that the two unipolar scales could be directly compared, while controlling for deviation from bipolar scale ratings. Responses to filler scales were not scored. The scoring procedure was as follows: each rating of a concept was assigned a number from 1 to 6, where 1 corresponded to the end of the scale representing the marked term (the term itself did not appear on the unmarked unipolar scales). The ratings of



each concept on each of the bipolar scales was subtracted from those on each of the two unipolar scales, yielding deviation scores. The absolute values of these scores were summed over the four concepts, producing (for each subject) a total deviation score for each End (marked *vs.* unmarked unipolar scale)  $\times$  Scale (four test scales) combination.

The resulting scores, averaged across subjects, appear in Table 1. These scores were subjected to an A (End)  $\times$  B (Scale) analysis of variance with repeated measures on both factors. This analysis indicated a significant End effect: ratings on marked unipolar scales deviated from bipolar scale ratings significantly more than did ratings on unmarked unipolar scales [ $F(1, 31) = 6.14, p \leq .05$ ]. While there was a significant Scale effect [ $F(3, 93) = 17.20, p \leq .001$ ], the End  $\times$  Scale interaction failed to reach statistical significance [ $F(3, 93) = 2.45$ , not significant], and so scale differences were not examined further.

### 3. Discussion

Results of Experiment I support the hypothesis that bipolar scale ratings resemble unmarked unipolar scale ratings more closely than the former resemble marked unipolar scale ratings. This finding is consistent with other studies using unipolar scales to examine symmetry (e.g., 1, 7), in that some evidence for asymmetry was obtained. (Examination of Table 1 suggests that the effect was not present for the scale long-short, although the interaction effect did not indicate this; in any case, in view of the polysemy of "short," which when not directly contrasted to "long" might be taken to refer to "not tall," such a lack of effect might be expected.) These results go beyond previous interpretations, in that the observed asymmetry was systematically related to lexical marking.

There seem to be at least two possible explanations for why lexical marking effects might arise in semantic differential scales. It is possible that the scales are structurally asymmetric: i.e., that the departure from symmetry results directly from characteristics of the scales. However, rating concepts on scales can be assumed to require cognitive activity on the part of the

TABLE 1  
MEAN DEVIATION SCORES, EXPERIMENT I

Scale	Unmarked form	Marked form
Good-Bad	.97	2.25
Long-Short	3.69	3.66
Clean-Dirty	1.88	2.19
Beautiful-Ugly	1.28	1.94



subject, who must after all read the concept, and read the end points of the scale. It is possible that in processing this information, there is a tendency to delete markers in the same way that subjects may delete markers in solving three-term series problems. In fact, one might even hypothesize that in order to rate a concept on a unipolar scale, the subject must transform that scale into a bipolar scale, a transformation which would require one more step for marked unipolar scales than for unmarked unipolar scales. Thus, a second explanation for the results of Experiment I might be that the cognitive processing demands of the rating process produced the lexical marking effect.

Experiment II was designed to replicate Experiment I, but was also intended to discriminate between the two hypotheses above. If the asymmetry in Experiment I were due to characteristics of the rating task as such, the magnitude of the effect might be expected to increase as the task became more demanding; on the other hand, if the asymmetry is inherent in the structure of the scales, increasing task difficulty should decrease the magnitude of the effect. Manipulation of such variables is relatively rare in paper-and-pencil research of this type, but is not unprecedented; for example, Trott and Jackson (19) found that an increase in the speed with which subjects were required to respond on such a task increased the amount of acquiescent response style revealed.

### C. EXPERIMENT II

#### 1. Method

*a. Subjects.* The subjects were 66 students enrolled in an introductory psychology course at Jackson Community College, Jackson, Michigan. Participation in the experiment was not required. Two subjects consistently marked ratings on one side of the pages and their data were discarded.

*b. Materials and Procedure.* There were three sections to each booklet, the first page being an instruction sheet. Subjects were given the same instructions as in Experiment I, except that they were told the experiment would have three parts, and that they would be informed of the time available before each part. Subjects were instructed to raise their hands immediately after completing each section, allowing the experimenter to time them. All subjects finished each part before the next part was begun.

The remainder of each booklet was composed of three sections, each consisting of seven pages, followed by a blank sheet indicating the end of each section. Concepts used in Experiment II were lake, lady, disease, feather, boulder, me, bird; scales were kind-cruel; strong-weak; active-passive; good-

bad; large-small; fast-slow; long-short; old-young; clean-dirty; and hot-cold. No filler scales were used in Experiment II.

In the first section, each concept appeared at the top of a page, followed by 10 bipolar scales. Side of page and order on page were counterbalanced across concepts. The seven pages in this section were assembled in random order. The next section consisted of the same concepts, but for each concept, all 10 scales were unipolar (half marked form, half unmarked form). The last section was the same as the second, except that for each scale, the complementary unipolar scale was substituted. In both sections, order of concepts was varied randomly. Half of the subjects in each group (see below) received booklets in the form above; the other half received booklets in which the order of sections two and three were reversed.

Subjects were run in two groups of 33 each (assigned randomly). Booklets were distributed, subjects read the instructions, and the experimenter answered any questions about procedure. On the experimenter's signal, the subjects began the first section. During this section, they were given as much time as they wished; in Group 1, the first subject finished in two minutes 10 seconds, which became the time limit for the speeded condition (see below; median time was about four minutes). Groups I and II varied only in the order in which they experienced the speeded condition. Group I received the speeded condition second and were given as much time as they wished during the third section (unspeeded condition). In Group II, the speeded condition was third. To enhance the demands placed on subjects during the speeded condition, in addition to the time constraint, a record described commercially as "distracting" was played ("Optimum Aviary," Syntonic Research, New York).

## 2. Results

Scoring procedure was the same as for Experiment I, except that deviation scores were summed over scales, as well as concepts, yielding (for each subject) total deviation scores for each End  $\times$  Speed (Speeded *vs.* Unspeeded condition) combination; see Table 2. These scores were subjected to an A (Group)  $\times$  B (End)  $\times$  C (Speed) analysis of variance, with repeated measures on B and C. In the summary which follows, effects not reported failed to reach statistical significance at the  $p \leq .05$  level.

The End effect was statistically significant: deviations from bipolar scale ratings were greater for marked unipolar scales than for unmarked unipolar scales [ $F(1, 62) = 4.93, p \leq .05$ ]. The main effect for Speed was also statistically significant [ $F(1, 62) = 5.12, p \leq .05$ ], suggesting that the

TABLE 2  
MEAN DEVIATION SCORES, EXPERIMENT II

Speed Cond.	Marked unipolar scales		Unmarked unipolar scales	
	Group I	Group II	Group I	Group II
Speeded	31.25	29.38	32.25	27.91
Not speeded	29.81	28.59	24.28	28.81

speed manipulation was effective in increasing the total deviation from bipolar scale ratings. However, two significant interactions involving the speed variable indicated that the manipulation was more effective in Group I than in Group II: i.e., an order effect was observed [ $\text{Group} \times \text{Speed}$ :  $F(1, 62) = 5.40$ ,  $p \leq .05$ ;  $\text{End} \times \text{Group} \times \text{Speed}$ :  $F(1, 62) = 4.26$ ,  $p \leq .05$ ]. This might well indicate a practice effect, since the speed manipulation was more effective in Group I, where the speeded condition was experienced second, than in Group II, where it was third. The hypothesized  $\text{Speed} \times \text{End}$  interaction failed to reach significance [ $F(1, 62) = 1.48$ , not significant], but in view of the order effect, Group I data were examined by means of a separate analysis. For Group I only, there was a significant  $\text{End} \times \text{Speed}$  interaction [ $F(1, 31) = 4.26$ ,  $p \leq .05$ ]: the magnitude of the difference between Ends was greater for the unspeeded condition than for the speeded condition. That is, increasing the difficulty of the task reduced the lexical marking effect.

### 3. Discussion

Results of Experiment II are consistent with those of Experiment I in that marked unipolar scale ratings deviated more from bipolar scale ratings than did unmarked unipolar scale ratings. This experiment involved different concepts and some different scales than did Experiment I. The lexical marking effect was attenuated by increased difficulty of the task, suggesting that the effect is not due to failure to process information accurately, but rather to a structural asymmetry in the scales themselves.

These experiments have implications for research with the semantic differential, and research relating to lexical marking. With regard to the question of scale symmetry in semantic differential research, the results obtained above provide additional evidence that the scales are not in fact symmetric. Specifically, departures from symmetry are systematically related to semantic structure of the terms defining the scales. As noted earlier, this does not mean that this is the only source of asymmetry.

The present study provides additional evidence for the existence of lexical

marking effects on cognitive processing tasks. In addition, the semantic differential technique used in the present study is potentially useful in the resolution of a pervasive problem in lexical marking research, the fact that lexical marking and connotative meaning (especially the Evaluation factor) are closely related. Thus, Hamilton and Deese (9) found that apparent lexical marking effects in word sorting can be confounded by the fact that most unmarked adjectives are evaluatively positive, while marked adjectives are negative. There are probably not enough scales or concepts in the experiments above to provide a definitive test, but if one were to use the unipolar scale technique with a number of scales which loaded on the Evaluative semantic differential factor to different degrees, and the asymmetry were due to evaluative differences between marked and unmarked terms, one would expect to obtain a Scale  $\times$  End interaction. That this interaction did not arise in Experiment I suggests that evaluation was not a confound, but this possibility requires further empirical examination.

## REFERENCES

1. ATWOOD, J. T., & FALKENBERG, S. D. The bipolarity of semantic space as related to concept complexity. *J. of Psychol.*, 1971, **79**, 119-122.
2. BENTLER, P. M. Semantic space is (approximately) bipolar. *J. of Psychol.*, 1969, **71**, 33-40.
3. CLARK, H. H. Linguistic processes in deductive reasoning. *Psychol. Rev.*, 1969, **76**, 387-404.
4. ———. The influence of language in solving three-term series problems. *J. Exper. Psychol.*, 1969, **82**, 205-215.
5. ———. Word associations and linguistic theory. In J. Lyons (Ed.), *New Horizons in Linguistics*. Harmondsworth, England: Penguin Books, 1970. Pp. 271-286.
6. CLARK, H. H., & CHASE, W. G. On the process of comparing sentences against pictures. *Cogn. Psychol.*, 1972, **3**, 472-517.
7. GREEN, R. F., & GOLDFRIED, M. R. On the bipolarity of semantic space. *Psychol. Monog.*, 1965, **79**, Whole No. 599.
8. GREENBERG, J. H. *Language Universals*. The Hague: Mouton, 1966.
9. HAMILTON, H. W., & DEESE, J. Does linguistic marking have a psychological correlate? *J. Verb. Learn. & Verb. Behav.*, 1971, **10**, 707-714.
10. HARRIS, R. J. Answering questions containing marked and unmarked adjectives and adverbs. *J. Exper. Psychol.*, 1973, **97**, 399-401.
11. HEISE, D. R. Some methodological issues in semantic differential research. *Psychol. Bull.*, 1969, **72**, 406-422.
12. HUTTENLOCHER, J., & HIGGINS, E. T. Adjectives, comparatives, and syllogisms. *Psychol. Rev.*, 1971, **78**, 487-504.
13. ———. On reasoning, congruence, and other matters. *Psychol. Rev.*, 1972, **79**, 420-427.
14. MALMSTROM, E. J., & FRENCH, G. M. Scale-symmetry and the semantic differential. *Amer. J. Psychol.*, 1963, **76**, 446-451.
15. MORDKOFF, A. M. An empirical test of the functional antonymy of semantic differential scales. *J. Verb. Learn. & Verb. Behav.*, 1963, **2**, 504-508.

16. ———. Functional vs. nominal antonymy in semantic differential scales. *Psychol. Rep.*, 1965, 16, 691-692.
17. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The Measurement of Meaning*. Urbana: Univ. Illinois Press, 1957.
18. TAYLOR, W. L., & KUMATA, H. Unpublished manuscript cited in C. E. Osgood, G. J. Suci, & P. H. Tannenbaum, *The Measurement of Meaning*. Urbana: Univ. Illinois Press, 1957. P. 153.
19. TROTT, D. M., & JACKSON, D. N. An experimental analysis of acquiescence. *J. Exper. Res. Personal.*, 1967, 2, 278-288.
20. WICKENS, D. D. Characteristics of word encoding. In A. W. Melton & E. Martin (Eds.), *Coding Processes in Human Memory*. New York: Wiley, 1972.

*Department of Psychology*  
*Michigan State University*  
*East Lansing, Michigan 48823*





## THE STEREOTYPES OF FOUR ETHNIC GROUPS\*<sup>1 2</sup>

*The University of Connecticut*

MENI KOSLOWSKY<sup>3</sup>

### SUMMARY

The possible interfering effects of social desirability (SD) on stereotype questionnaires were investigated. By controlling SD, it was assumed subjects would more likely respond to the qualitative meaning of an adjective, and unique stereotypes would result for each nationality. The Katz and Braly adjectives were presented to 800 students in a forced choice format. Two different questionnaires—one controlling for SD, the other for applicability or frequency (but not for SD)—were analyzed for Israelis, Jews, Japanese, and Germans. Results showed that when SD was controlled, negative as well as positive adjectives were chosen to describe the groups. When SD was not controlled, frequency of choice and SD correlated substantially. However, except when Japanese were in the comparison, stereotypes overlapped each other.

### A. INTRODUCTION

Stereotype research, which dates back 40 years (5), has undergone a revival in the past decade.<sup>4</sup> Several of the crucial studies in this area have come from researchers at Princeton (3, 4, 5). Using the 84 adjectives introduced by Katz and Braly in 1933, Karlins, Coffman, and Walters (4) recently reported high agreement in the stereotypes used to describe 10 national groups. Karlins *et al.* also obtained mean estimates of favorableness for each of the 84 traits and found that within most stereotypes positive values outweighed negative values.

Two possible explanations for the last results were offered by the authors:

\* Received in the Editorial Office, on August 10, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This article was based on the author's doctoral dissertation submitted to Teachers College, Columbia University, New York. The author was supported by Educational Research Traineeship No. 20-2024-081 (082).

<sup>2</sup> Reprints may be obtained from the author at the address shown at the end of this article.

<sup>3</sup> The author acknowledges his appreciation to R. L. Thorndike for his advice in the study reported here.

<sup>4</sup> For an excellent review summarizing the present conceptualizations of stereotypes in psychology, see Brigham (2).

(a) the 10 groups possessed genuinely favorable images which were represented by the responses in the three studies, or (b) there is a tendency for college students, particularly in today's liberal atmosphere on most campuses, to respond with positive words when describing ethnic groups, even if these are not the actual attitudes of the subjects. Consequently, very few negative adjectives would appear in any group's stereotype. Many more might be applicable, but if social desirability (SD) responding was influencing the results, they would not be chosen. Reducing the possible interfering effects of SD in stereotype questionnaires was a major concern of this study.

This reduction in the influence of SD should also result in unique stereotypes for each of the groups. As subjects focus on the relevance, not the SD value, of an adjective for a particular group, different response patterns for each of the groups can be expected to emerge.

By using a verbal technique dubbed the "bogus pipeline," Sigall and Page (6) showed that social desirability can influence results in a stereotype investigation. A major drawback which limited the generalizability of the results was the fact that Sigall and Page used verbal responses to gauge stereotyping, whereas most previous researchers had used written responses. The present study used the traditional checklist method but in a forced choice format (1, 7) to control for SD.

Essentially, this technique requires respondents to choose between two or more descriptive terms or phrases that have been matched as closely as possible on a particular dimension. The two dimensions relevant to the Katz and Braly adjectives are SD and applicability. The former can be ascertained by having subjects rate the adjectives on some scale of favorability, whereas the latter can be measured by determining the frequency with which particular adjectives have been used in describing the groups. The alternatives within a group are matched as closely as possible on the relevant variable. The grouped terms may all be desirable or all undesirable, all applicable or all inapplicable.

If SD has been effectively controlled, subjects will attend less to the SD value of a word and more to its relevance for a particular nationality. With this premise, the following hypotheses emerge:

1. When social desirability is controlled, the frequency with which certain adjectives are chosen to describe a group will differ significantly from random responding, and this will be true for unfavorable, as well as favorable, adjectives. When social desirability is allowed to vary, frequency of selection of adjectives will be related to social desirability.

2. When social desirability is controlled, there will be a difference in the content of the stereotypes attributed to different groups.

## B. METHOD

### 1. Instruments

The present study used two parallel questionnaires, one controlled for SD and the other for applicability, to examine the stereotypes concerning four groups: Israelis, Jews, Japanese, and Germans. The questionnaires were derived from the 84 adjectives used by Katz and Braly (5). Karlines and his co-workers obtained ratings from 100 subjects on favorability or desirability for each of the adjectives. By using a five-point,  $-2$  to  $+2$ , scale, with  $-2$  representing most unfavorable and  $+2$  most favorable, and calculating a mean rating, they were able to get a relative measure of the SD value for each word. The words were then ordered from highest to lowest.

For the purposes of this study, it was decided to leave out the three adjectives rated most undesirable in the Karlines *et al.* study. This was done for two purposes. In order that a square block design (described below) be used to control for extraneous variables, it was necessary to have a number which has as its square root an integer. Leaving out any three adjectives would not adversely affect the major objectives of this study. Consequently, it was decided to eliminate the three adjectives with the lowest SD values—"rude," "deceitful," and "cruel." It was felt that their elimination would make the rating task that much more bearable.

The present study employed a nine by nine square block technique in order to permit the examination of a particular dimension while controlling the influence of a second dimension. In order to control for SD, nine groups of nine words were formed. These divisions were based on the mean SD values given by the raters in the Karlines *et al.* study (4, p. 10). All words were first ordered from highest to lowest according to their mean SD values. The word "brilliant" with a mean value of  $+1.70$  is the first word, while the word "treacherous" with a mean value of  $-1.65$  is the last word. The 81 words were then divided into nine groups. The first group contains words 1 through 9 in the SD ordering ("brilliant" through "faithful"), the second group contains words 10 through 18 ("sportsmanlike" through "progressive"), until the ninth list, which contains words 72 through 81 ("revengeful" through "treacherous") in the SD ordering. Each group of nine words contains adjectives with similar SD values. The SD values within groups range from .20 in the second group to .47 for the first group.

The adjectives within each group, as well as the groups themselves, were randomized before presentation. The identical scheme was used for each nationality. The instructions on each questionnaire required the rater to choose the three words from each set (group of nine adjectives) that he thought were given as most characteristic for that nationality. The instructions were worded indirectly because it was felt that the direct question, "What traits do you think of as most characteristic . . . ?" would offend too many people and result in many more questionnaires going unanswered.

On the basis of the data gathered from this questionnaire, the nine words from each set were ordered to reflect their relative standing in applicability, as measured by the frequency with which the adjectives were chosen in the SD controlled questionnaire. A new nine by nine square block was then formed for each nationality. Each of these new groups contained nine words with the same rank in frequency. The first group contained the word most often chosen in each of the nine sets from the SD controlled questionnaire; the second group contained the second most often chosen word in each set from the SD controlled questionnaire and so on, until the ninth group, which contained the words chosen least often from the SD controlled sets. The degree of similarity within the sets can best be illustrated with a typical example from the questionnaires: In Set 6 of the applicability controlled questionnaire for Japanese, the range in frequency is from zero ("kind," "lazy," and "stupid") to three ("conservative").

Since the rank in frequency for the adjectives usually varied amongst nationalities, the content of the nine new sets was different for each ethnic group. These nine sets formed the applicability controlled questionnaire.

Before presentation, the adjectives within each group, as well as the groups themselves, were randomized. The instructions, as well as the structure, of the applicability controlled questionnaires were the same as those of the social desirability controlled questionnaires.

By the use of these two types of questionnaires, two separate conditions were set up. The first questionnaire, which controlled for SD, permitted applicability to vary. The second questionnaire, which controlled for applicability, required SD to vary. Analysis of the results was done separately for each condition.

## 2. *Raters*

Altogether 800 questionnaires were used, 400 for each different form of the questionnaire. To achieve this total, over 1200 questionnaires were handed out. (Wastage occurred for several reasons: refusal to answer the question-



naire, not following instructions correctly, failure to return the questionnaire, etc.) Subjects were undergraduate or graduate students in New York City universities. For the SD controlled questionnaire, there were 214 males and 186 females responding; the popularity controlled questionnaire had 170 male and 230 female respondents. Equality in numbers of each sex was the preferred distribution of questionnaires. However, the nature of random sampling made it difficult to achieve this goal.

### 3. Procedure

The different questionnaires were handed out randomly in the several universities comprising the study. A student not wishing to respond to the questionnaire he was given was not permitted to choose another one. If a student agreed to participate in the study, he was told to follow the written instructions on the questionnaires. Raters were required to follow these directions on their own, as no further communication between examiner and rater was allowed. Although no specific time limit was given, any questionnaire not completed in about 15 minutes was usually left for the respondent and not analyzed in the study.

One hundred different subjects (raters) responded to each nationality under each of the two conditions (questionnaires). They were required to select three adjectives from each set for a total of 300 responses per set for each nationality. During the scoring procedure, any set that contained more or less than three responses disqualified the entire questionnaire.

## C. RESULTS

### 1. Results Relating to Hypothesis 1

To test whether the results for the SD controlled questionnaire were significant, a  $\chi^2$  goodness of fit test was used. The chi square was determined by adding up the squared differences between the observed and expected frequencies divided by 33.33, the expected frequency for each word. Rounding off was done in such a way that the obtained  $\chi^2$  would always be lowest. Results were significant for each set beyond the .001 level (Table 1).

The  $\chi^2$  model does not exactly fit here because the responses are not all independent, as required by the model. The fact that each subject made three choices introduces dependency and works against obtaining significance as it tends to "flatten out" the distribution. However, the obtained values of  $\chi^2$  are so large that this deficiency should not make much difference.<sup>5</sup>

<sup>5</sup> Professors Robert L. Thorndike, a measurement psychologist, and Richard H. Lindeman, a statistician, advised in this study and concurred with the statement.

TABLE 1  
 $\chi^2$  AND CORRELATIONS BETWEEN RANK IN SOCIAL DESIRABILITY AND RANK  
 IN FREQUENCY FOR EACH SD CONTROLLED SET

Sets	Israelis		Jews		Japanese		Germans	
	$\chi^2$	R	$\chi^2$	R	$\chi^2$	R	$\chi^2$	R
1	124.74	.07	76.71	.40	48.86	-.08	142.74	.00
2	216.49	-.01	151.97	-.40	79.65	.29	132.52	.19
3	126.77	-.26	116.29	-.52	76.06	-.63	99.49	.00
4	95.44	-.28	172.93	-.12	222.56	-.17	98.60	-.29
5	175.02	-.15	118.54	-.10	28.21	.12	147.55	-.22
6	83.93	.07	117.78	.10	68.00	.05	67.75	.28
7	72.42	.12	76.71	-.05	46.81	.30	64.17	.00
8	95.86	.50	105.02	.52	70.11	.65	70.91	.57
9	204.15	.40	106.89	.22	98.86	.35	154.55	.35
Mean		.05		.01		.10		.10

Note: All chi squares are significant at  $p < .001$ .

The SD controlled results reveal some of the opinions students in this sample had of these four groups.<sup>6</sup> Many of the obtained frequencies deviate substantially from 33.33, the expected total frequency for each word in each group. Adjectives such as "generous" or "witty" do not seem to be very characteristic of any group, whereas 85% of the subjects found Japanese "loyal to family ties" and 82% found Israelis "arrogant."

Since the adjectives in each SD controlled set had a range of values, it was necessary to examine the relationship between frequency of choice and SD within each set. In order to make sure that responding within a subset was not a function of SD, a rank order correlation between frequency rank and rank in SD was calculated for each set of nine words. The average obtained  $R$  across all sets and groups was .065. This result would indicate that distortion on the basis of SD alone had been effectively controlled for the group as a whole.

This does not hold true for the questionnaire where SD varied within each set. As hypothesized, stereotypes in this situation correlated rather substantially with SD. The average correlation between frequency rank and rank in SD was .54 for Germans, .86 for Israelis, .67 for Jews, and .78 for Japanese. The average of all the rank order correlations was approximately .71.

## 2. Results Relating to Hypothesis 2

It was predicted that the four ethnic groups would differ in the types of adjectives attributed to them. To test this hypothesis, two statistics were

<sup>6</sup> For a complete breakdown of the adjectives chosen for each group, order NAPS Document No. 02200 from Microfiche Publications, 305 East 46th Street, New York, New York 10017; remit \$1.50 for microfiche or \$5.00 for photocopies.

used:  $\chi^2$  and rank order correlations. The  $\chi^2$  was calculated for each list for all possible combinations of groups (Table 2). A substantial number of the tests are significant at the .001 level, particularly when Japanese are compared to any other group. Such a stringent level was used because of the many  $\chi^2$  tests that were carried out and because the requirement of independence was not met.

Although it has been demonstrated that there were significant differences between national groups in the frequency with which certain words were chosen to describe them, it must also be recognized that there was a general tendency for the several national portraits to be similar. Except when the Japanese are in the comparison, the adjectives seem to rank pretty much in the same order for each group. The mean  $R$  for Israelis *versus* Jews is .83 and for Israelis *versus* Germans, .85. The mean  $R$  for Jews *versus* Germans is .67. The Japanese, however, show an average correlation of .49 with Israelis, .51 with Jews, and .53 with Germans.

The second hypothesis was also tested by using rank order correlations. When SD was controlled, these correlations averaged out to .65 (see Table 2). A relatively high correlation such as this seems to indicate that although some distinctions existed between groups, a great amount of overlap in stereotype was present.

#### D. DISCUSSION

The present study was primarily concerned with two major issues: (a) the degree to which stereotypes, such as those determined in past questionnaires, are a function only of positive reactions to these nationalities and the degree to which there is also a sizable component that represents content apart from the general positive reaction; (b) the differences, if any, that exist in the stereotypes of Jews, Israelis, Germans, and Japanese.

Results showed the following: (a) subjects permitted to respond on the basis of SD tended to do so but definite patterns of perception still existed when SD was controlled; (b) the adjectives comprising the stereotypes of these four groups were more alike than they were different. Differences that existed were principally between Japanese and the other groups.

No formal evaluation of the degree of contact was done in the present study, and one can raise the query of how contact, or the lack of it, influenced the uniformity of stereotypes. Similarly, the role of the mass media was not analyzed. The investigation of these intervening variables seems to be an interesting line of inquiry for the future.

In studies of the type discussed in this paper, the method itself often

TABLE 2  
 $\chi^2$  AND CORRELATIONS FOR RANK IN FREQUENCY FOR EACH SET BETWEEN EACH ETHNIC GROUP  
 ON THE SD CONTROLLED QUESTIONNAIRE

Sets	Jews			Israelis versus			Germans			Japanese			Jews versus			Germans			Japanese versus		
	$\chi^2$	$R$		$\chi^2$	$R$		$\chi^2$	$R$		$\chi^2$	$R$		$\chi^2$	$R$		$\chi^2$	$R$		$\chi^2$	$R$	
1	14.19	.81		38.22*	.45		7.45	.96		34.26*	.56		20.84	.77		25.09	.61		20.84	.77	
2	33.77*	.78		89.87*	.40		15.95	.90		81.63*	.28		61.41*	.65		46.55*	.58		61.41*	.65	
3	12.30	.92		42.12*	.63		30.79*	.58		37.28*	.73		51.65*	.42		35.14*	.51		51.65*	.42	
4	13.70	.93		20.89	.89		17.14	.85		16.24	.86		33.91*	.79		27.37*	.95		33.91*	.79	
5	34.08*	.79		96.17*	.36		16.95	.92		69.96*	.15		42.84*	.68		70.09*	.45		42.84*	.68	
6	26.84*	.87		53.44*	.08		9.59	.89		40.73*	.39		17.20	.86		40.89*	.42		17.20	.86	
7	12.52	.81		49.59*	.09		6.35	.87		49.61*	.40		19.50	.55		60.21*	.33		19.50	.55	
8	9.98	.87		48.68*	.57		22.25	.75		49.23*	.40		28.45*	.64		37.30*	.97		28.45*	.64	
9	37.06*	.70		22.41	.90		12.19	.90		40.24*	.79		56.40*	.65		11.16	.97		56.40*	.65	
Mean		.83			.49			.85			.51			.67			.53			.67	

\*  $p < .001$ .

influences the results. Consequently, comparison of findings among different methodologies is quite hazardous. As stated previously, verbal techniques, as used by Sigall and Page, can lead to different findings than written scales. Similarly, results from written scales are often a function of the type of questionnaire used. The model for stereotype research introduced by Katz and Braly required subjects to respond to only five out of a possible 84 adjectives. The present study required three responses from each set of nine, for a total of 27 adjectives chosen out of a possible 81. The task is obviously much more difficult in this latter case.

Furthermore, a forced choice format was used in the present study. One must be very careful about the conclusions drawn from such a measure, often referred to as ipsative rather than the usual normative (1). Many of the more common statistical procedures are not valid here. One can make statements about the relative frequencies with which particular adjectives within lists were chosen and compare these results to those of another group. However, one cannot accurately conclude anything about the absolute frequency with which particular adjectives are used in describing a group.

The above is yet another reason why care must be taken in making any generalizations from the present research. The significant findings from this study are more methodological than substantive. The forced choice technique has helped show that SD can influence the quality of adjectives comprising a national stereotype and must be controlled before meaningful conclusions can be drawn.

#### REFERENCES

1. ANASTASI, A. *Psychological Testing* (3rd ed.). New York: Macmillan, 1968.
2. BRIGHAM, J. C. Ethnic stereotypes. *Psychol. Bull.* 1971, **76**, 15-38.
3. GILBERT, G. M. Stereotype persistence and change among college students. *J. Abn. & Soc. Psychol.* 1951, **46**, 245-254.
4. KARLINS, M., COFFMAN, T. L., & WALTERS, G. On the fading of social stereotypes: Studies in three generations of college students. *J. Personal. & Soc. Psychol.* 1969, **13**, 1-16.
5. KATZ, D., & BRALY, K. Racial stereotypes in one hundred college students. *J. Abn. & Soc. Psychol.* 1933, **28**, 280-290.
6. SIGALL, H., & PAGE, R. Current stereotypes: A little fading, a little faking. *J. Personal. & Soc. Psychol.*, 1971, **18**, 247-255.
7. THORNDIKE, R. L., & HAGEN, E. *Measurement and Evaluation in Psychology and Education* (3rd ed.). New York: Wiley, 1969.

*Department of Behavioral Sciences and Community Health  
School of Dental Medicine  
The University of Connecticut  
Farmington, Connecticut 06032*





## THE RELATIVE EFFECTIVENESS OF TWO METHODS OF PRESENTING VISUALIZED INSTRUCTION\*

*University Division of Instructional Services, The Pennsylvania State University*

---

FRANCIS M. DWYER

---

### SUMMARY

The purpose of this study was to investigate the relative effectiveness of two methods of presenting visualized instruction to Ss and to determine the effectiveness of different types of visuals in facilitating S achievement of different educational objectives. Five hundred twenty Ss participated in the study; each received a pretest, participated in his respective instructional presentation, and received four criterial tests. Analysis indicated that (a) the method of presenting visualized instruction determines the type of visualization most effective in facilitating S achievement of different educational objectives; (b) for facilitating S achievement of certain objectives, color in visuals is an important instructional variable; and (c) not all visuals are equally effective in facilitating S achievement of different educational objectives.

---

### A. INTRODUCTION

Considerable attention is currently being focused on the use of visual materials to complement regular classroom instruction. Although previous research (3) has established that the use of visuals used to complement instruction improves S achievement, little experimental evidence is available indicating which types of visual illustrations are most effective in facilitating S achievement of specific types of educational objectives when different instructional formats are employed to present information to Ss. Thus, the purposes of this study were (a) to measure the relative effectiveness with which two methods of presenting visualized instruction (slide/audiotaped and programmed instruction) facilitate S achievement on criterial tests designed to measure different educational objectives; (b) to determine whether color in illustrations, presented via different instructional formats, is an important instructional variable in increasing S achievement on different

---

\* Received in the Editorial Office, on August 13, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

critical measures; and (c) to identify the type of visual most efficient in facilitating *S* achievement on critical tests measuring different educational objectives.

## B. METHOD

### 1. *Treatment Groups*

Five hundred twenty *Ss* at The Pennsylvania State University were randomly assigned to one of 18 treatment groups. Nine groups received the slide/audiotaped instruction and nine received the programmed instruction (See Table 1). A description of the procedures involved in developing the slide/audiotaped and programmed treatments has been presented in previous articles reported in this journal (1, 2). The content material used in this study concerned the human heart, its parts, and functions.

TABLE 1  
S ASSIGNMENT TO INSTRUCTIONAL GROUPS

Instructional format	Group								
	I	II	III	IV	V	VI	VII	VIII	IX
Programmed	31	29	28	31	26	29	28	30	27
Slide/audiotape	28	27	32	30	29	29	27	30	29

Information transmitted via the printed and oral channels was held constant for each presentation and the amount of realistic detail contained within the visual illustrations was varied. *Ss* in Group I in each presentational format (slide/audiotaped and programmed) received no visual illustrations but viewed printed words; Group II viewed simple line illustrations (black and white); Group III viewed simple line illustrations (colored); Group IV viewed detailed, shaded drawings (black and white); Group V viewed detailed, shaded drawings (colored); Group VI viewed photographs of a heart model (black and white); Group VII viewed photographs of a heart model (colored); Group VIII viewed realistic heart photographs (black and white); and Group IX viewed realistic heart photographs (colored).

### 2. *Procedure and Critical Measures*

Each *S* in each treatment group received the Otis Quick Scoring Mental Ability Test (Form Fm) as a pretest, participated in his respective instructional presentation, and then received four individual critical tests. Scores

received on these tests were combined into a 78-item total criterial test. The objective of each test was as follows: (a) drawing test—to evaluate learning of specific locations of the various patterns, structures, and positions of the parts of the heart; (b) identification test—to measure transfer of learning; i.e., the ability to identify numbered parts on a diagram of the heart from information presented via the instructional presentation; (c) terminology test—to evaluate knowledge of referents for specific symbols; (d) comprehension test—to measure understanding of the heart, its parts, and its internal operations; and (e) the total criterial test—to measure the Ss total understanding of the concepts presented. The Kuder-Richardson Formula 21 reliability coefficients for the five criterial measures were: (a) drawing test .71; (b) identification test .83; (c) terminology test .75; (d) comprehension test .79; and (e) total criterial test .94.

### C. DESIGN

The Bartlett test for homogeneity of variance and ANOV conducted on the IQ scores Ss achieved on the Otis Quick Scoring Mental Ability Test did not reach the critical value for an .05 level test, indicating that the treatment groups were drawn randomly from populations with common variance. A *t*-test analysis was conducted on achievement scores obtained by Ss receiving corresponding instruction (Group I programmed *versus* Group I slide/audiotape) presented via the different instructional formats.

### D. RESULTS

Table 2 shows where significant differences in achievement (.05 level) occurred between Ss receiving the two different methods of presentation.

TABLE 2  
METHOD OF PRESENTATION MOST EFFECTIVE IN FACILITATING ACHIEVEMENT  
ON EACH CRITERIAL MEASURE

Verbal criterial measures	Instructional treatments							
	Simple line drawing		Detailed drawing		Heart model		Realistic photographs	
	B & W	Color	B & W	Color	B & W	Color	B & W	Color
Drawing		S > P					P > S	P > S
Identification	P > S			S > P				P > S
Terminology					P > S			P > S
Comprehension					P > S			P > S
Total criterial					P > S			P > S

Note: P = Programmed instruction; S = Slide/Audiotaped Instruction; B & W = black and white.

The blank areas indicate that significant differences in achievement did not occur.

### E. DISCUSSION AND SIGNIFICANCE

Ss receiving the programmed instruction format achieved significantly higher scores when they received the more realistic instructional treatments. One explanation for the relative success of the more realistic illustrations presented in the programmed instruction format might be that, because instruction was internally paced, students were able to interact with the more realistic visualization for as long as necessary in order to comprehend the information being presented. These results seem to substantiate the current trend towards visualized self-paced instruction as being an effective instructional technique.

A further interpretation of the data also indicated that (a) all types of visuals are not equally effective in facilitating S achievement of different educational objectives; (b) for certain objectives the use of visualization to complement instruction does not improve S achievement; (c) the type of visualization most effective in facilitating S achievement of a specific objective is dependent on the method of instruction employed to present the information to the S; and (d) the use of color in certain types of visuals to facilitate S achievement of specific objectives is an important instructional variable.

### REFERENCES

1. DWYER, F. M. The instructional effect of motion in varied visual illustrations. *J. of Psychol.*, 1969, 73, 167-172.
2. ———. Effect of knowledge of objectives on visualized instruction. *J. of Psychol.*, 1971, 77, 219-221.
3. ———. A Guide for Improving Visualized Instruction. State College, Pa.: Learning Services, 1972.

*University Division of Instructional Services  
The Pennsylvania State University  
Instructional Services Building  
University Park, Pennsylvania 16802*



## MASSSED *VERSUS* SPACED SESSIONS IN SYSTEMATIC DESENSITIZATION\*<sup>1</sup>

*College of St. Teresa*

---

JOHN O. KLING AND ELLEN M. MINOGUE

---

### SUMMARY

Research has indicated that massed systematic desensitization can be effective. There is no definite evidence, however, of the superiority of massed or spaced therapy. The present study investigated the effects of massed and distributed systematic desensitization upon a spider phobia. Eighteen Ss were given a fear hierarchy before therapy and then after therapy. Analysis of the results suggests that massed and distributed sessions do not differ significantly in effectiveness, which supports Suinn and Hall (5).

---

### A. INTRODUCTION

Therapy sessions in systematic desensitization typically have relied on two or three meetings a week. It is possible, however, to "mass" these sessions into a single meeting per week (3). At the same time, the research on spacing of sessions in systematic desensitization has provided no definitive evidence of the superiority of either massed or spaced sessions in therapy (3, 6). Methodologically, much of the research on massed-spaced sessions has left something to be desired. Studies have lacked control groups (3, 4, 6) and have not equated the amount of time in therapy for the experimental groups (6). The present study, using a control group and equating time in therapy for both massed and distributed sessions, will investigate the effects of massed and distributed systematic desensitization upon spider phobia.

### B. METHOD

#### 1. *Subjects*

Ss were 18 female students enrolled in the College of St. Teresa who volunteered to participate in a study requiring a fear of spiders. Each S

---

\* Received in the Editorial Office, August 16, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This paper is based on an independent study presented by the junior author to the Psychology Department, College of Saint Teresa, for the B. A. degree.

was randomly assigned to one of the three groups: spaced therapy (ST), massed therapy (MT), or control receiving no therapy (NT). A one-way analysis of variance on the pretest therapy ratings indicated that the groups were equal ( $F = .982, df = 2, 15, p > .05$ ).

## 2. Procedure

By volunteering for the study each *S* had admitted at least a mild fear of spiders. Prior to the beginning of therapy all *Ss* rated a fear hierarchy on a scale of 7 to 10. The ratings reflected the amount of anxiety that each of the proposed situations would cause the subject.

Both MT and ST were desensitized in groups by listening to a tape by the therapist. In addition, all three groups received relaxation training with Wolpe's technique (7). The ST group met 30 minutes a session for eight sessions; the MT group met for four hours all one day; the NT group met only twice—once for a pretest measure of anxiety and again for relaxation training and a posttest of anxiety. Fear situations on the hierarchy were presented to both experimental groups at one minute intervals. The two groups were equal in the number of fear presentations, time elapsed between each fear situation on the hierarchy, time spent in therapy, and time from the last therapy session to the posttest.

## 3. Results

A two-way analysis of variance with repeated measures on the subject variable was conducted upon the pretest and posttest fear hierarchy data. The analysis indicated that a significant difference existed between therapies ( $F = 3.05, df = 2, 15, p < .01$ ); also, there was a significant difference between pre- and posttest therapies ( $F = 40.58, df = 1, 18, p < .01$ ); and a significant interaction between therapies and pre-post therapies ( $F = 8.15, df = 2, 15, p < .01$ ). Graphing the interaction indicated that the significance was caused by the control group which did not change as much from pretherapy to posttherapy as did the two experimental groups.

Since differences existed between therapies, an internal test was conducted on the group means. A Neuman-Keuls test indicated that spaced and massed therapy were not significantly different. However, both differed significantly from the control group (NT).

## 4. Discussion

The data suggest that massed and distributed sessions of systematic desensitization do not significantly differ in their effectiveness. This conclu-

sion supports the results of Suinn and Hall (6). However, the present results disagree with Ramsay, Barends, Breuker, and Krusemen (3) who found that massed was superior to spaced. In accounting for these differences one must look at the different methodologies employed in the present study and Ramsay *et al.* The present study massed therapy into one four-hour session, while Ramsay *et al.* massed therapy into two 40-minute sessions separated by four days. Consequently, in the present study there was no intertherapy time lapse for the massed group. Second, for the ST group the present study employed 30-minute sessions a day as opposed to 20 minute sessions used in Ramsay *et al.* Thus, the present study appears to have massed therapy to a greater extent than Ramsay *et al.* (3).

The present data support the results of several studies which have shown the effectiveness of group therapy settings for systematic desensitization (1, 2) and the effectiveness of massing therapy into one session instead of the traditional three times a week (4, 5).

#### REFERENCES

1. LAZARUS, A. A. Group therapy of phobic disorders by systematic desensitization. *J. Abn. & Soc. Psychol.*, 1961, **66**, 519-525.
2. PAUL, G. L. *Insight versus Desensitization in Psychotherapy*. Stanford, Calif.: Mayfield, 1966.
3. RAMSAY, R. W., BARENDT, J., BREUKER, J., & KRUSEMEN, A. Massed versus spaced desensitization of fear. *Behav. Res. & Ther.*, 1966, **4**, 205-207.
4. ROBINSON, C., & SUINN, R. M. Group desensitization of a phobia in massed sessions. *Behav. Res. & Ther.*, 1969, **7**, 319-321.
5. SUINN, R. M. Case histories and shorter communications: Short-term desensitization therapy. *Behav. Res. & Ther.*, 1970, **8**, 181-184.
6. SUINN, R. M., & HALL, R. Marathon desensitization groups: An innovative technique. *Behav. Res. & Ther.*, 1970, **8**, 97-98.
7. WOLPE, J. *The Practice of Behavior Therapy*. Elmford, N. Y.: Pergamon, 1969.

Department of Psychology  
College of Saint Teresa  
Winona, Minnesota 55987



## MANAGERIAL ATTITUDES TOWARD HIRING EX-CONVICTS\* 1

*Faculty of Business Administration and Commerce, The University of Alberta, Canada*

---

ALBERT N. B. NEDD

---

### SUMMARY

It is generally assumed that the attitudes of managers are somehow determinant of the policies of organizations toward the recruitment of ex-convicts. This study demonstrates that the recruiting policies of organizations are independent of the attitudes of managers. Attitudes of managers toward hiring ex-convicts were found to be negatively related to the age of the organization and to have a mixed relationship to the age, level of education, and length of service of managers.

---

### A. INTRODUCTION

The purpose of this study was to ascertain what personal or organizational factors are associated with positive or negative managerial attitudes toward hiring ex-convicts. Ex-convicts were defined as persons who had served a term of imprisonment following conviction for a criminal offence. It was assumed that an important goal of our penal systems was the rehabilitation of prisoners. It was also assumed that the systematic absorption of discharged prisoners into the labor force, on an equal basis with nonoffenders, is essential to their effective rehabilitation.

The research design of the study was based on the proposition that the attitudes of managers have a bearing on the employment opportunities that are available to ex-convicts. Attitudes are generally defined in the literature as the constellation of situationally and culturally defined beliefs, values, and norms pertaining to a specific object in a person's cognitive world (1, 6). Thus, it was assumed that a manager's attitudes towards ex-convicts would determine whether the policies of his organization facilitated or hindered the employment of ex-convicts. For example, if a manager believed that criminal conduct was akin to drug addiction, in that a convict would normally be expected to persist in his criminal activities, it was felt that this belief would

---

\* Received in the Editorial Office on August 20, 1973, and published immediately at Provincetown, Massachusetts. Copyright by the Journal Press.

<sup>1</sup> This research project was partially financed by a grant from the Solicitor-General's Department, Ottawa, Canada.



operate to bias him against the employment of ex-convicts. On the other hand, if he believed that employers should assist discharged prisoners to readjust to society, he might be motivated to facilitate their employment by his organization.

The literature suggests that the rehabilitation of criminals is best accomplished by their resocialization, a process which results in their becoming motivated to accept legitimate patterns of behavior and to avoid illegal conduct. However, whether resocialization of a criminal in fact occurs will depend on the extent to which the opportunity structure of society operates to alleviate the strains that motivate him towards deviant behavior (5). Merton observed that "Owing to their objectively disadvantaged position in the group, as well as to distinctive personality configurations, some individuals are subjected more than others to the strains arising from the discrepancy between cultural goals and effective access to their realization. They are, consequently, more vulnerable to deviant behavior" (2, pp. 233-234). This, he argues, leads to the development of anomie defined as "a breakdown in the cultural structure, occurring particularly when there is an acute disjunction between cultural norms and goals and the socially structured capacities of members of the [society] to act in accord with them" (2, p. 216). The conceptual framework of this study assumed that such a disjunction occurs when the beliefs and values of managers operate to deny employment opportunities to ex-convicts. Difficulty encountered by ex-convicts in returning to the mainstream of society tends to drive them to return to their criminal activities. As Rubington and Weinberg observe, the fact that recidivism is occasioned by the operation of the social system against ex-convicts might be evidence that the social system is working as intended rather than malfunctioning (5). The point being made here is that while cultural goals are overemphasized, the social structure operates to prevent certain groups from obtaining them through legitimate means.

In order to identify some of the personal and organizational factors that are associated with positive or negative managerial attitudes towards employing ex-convicts, the following hypotheses were investigated in this study:

*Hypothesis 1:* There will be a positive correlation between the attitudes of managers and the policies of their organizations toward employing ex-convicts.

*Hypothesis 2:* The greater the personal sense of security of the manager the more positive will be his attitudes toward hiring ex-convicts.

It was assumed that a manager's personal sense of security will vary

directly with his age, level of formal education, and length of service with his organization. Hence, Hypothesis 2 might be restated as follows:

*Hypothesis 2(a):* the favorability of managers' attitudes toward the employment of ex-convicts will vary directly with their age, level of education, and length of service with their organizations.

*Hypothesis 3:* The greater the stability and success of their organization the more positive will be managers' attitudes toward hiring ex-convicts.

For the purposes of this study we concluded that organizational success and stability would be adequately reflected by the size of the organization (as determined by the number of persons it employed) and the length of time it had existed. Hence, Hypothesis 3 might be restated as follows:

*Hypothesis 3(a):* The larger the organizations and the longer they have been in existence, the more positive will be managers' attitudes towards hiring ex-convicts.

## B. RESEARCH METHOD

Subjects comprised managers from a random sample of organizations in a large Canadian city. The organizations were selected from the records of the local Chamber of Commerce and from lists published by the Provincial Government. The sample was stratified according to such factors as organization size and whether they were publicly owned or private institutions. Questionnaires were mailed to 350 organizations, and responses were solicited from managers who had functional responsibility for the formulation and implementation of employment policies. These were either personnel specialists or general managers, depending on the size of the organizations. Only one reply was accepted from each organization. A total of 217 managers returned acceptable questionnaires, and these, therefore, constituted the subjects in the study.

Attitudes toward the employment of ex-convicts were measured by subjects' responses to an item in the questionnaire constituting a six-item scale of orientation towards the employment of ex-convicts. Responses were scored 1 to 6, with lower scores indicating negative attitude towards hiring ex-convicts. Subjects were also requested to respond to a five-item scale indicating the degree of favorability of the official policies of their organizations towards hiring ex-convicts. Here, too, a low score also indicated a negative policy towards hiring ex-convicts.

A standard statistical package was employed to calculate Pearson correlation coefficients for variables used to test the hypotheses. Significance tests

for each coefficient were performed with use of Student's  $t$  with  $n-2$  degrees of freedom (4).

### C. RESULTS AND DISCUSSION

The hypotheses were tested for the sample as a whole. The sample was also subdivided into the following categories:

1. By type of organizations: Subjects were classified as being employed by (a) publicly owned organizations, (b) privately owned primary or manufacturing organizations, and (c) privately owned service organizations.

2. According to the main skill level of employee recruited by their organizations during the preceding two years. For this purpose, subjects were classified as employed by organizations recruiting mainly (a) unskilled or semiskilled labor, (b) skilled labor, (c) clerical personnel, or (d) professional and technical personnel.

3. By nature of their jobs: In this case subjects were classified either as (a) personnel specialists, or as (b) general managers.

The hypotheses were tested separately for each classification.

#### 1. Hypothesis 1

Table 1 (first column) indicates that the data for the sample as a whole did not support Hypothesis 1. The policies of organizations towards hiring ex-convicts were independent of the attitudes of their managers. This finding suggests that efforts to change the attitudes of managers will have relatively little impact on the rate of employment of ex-convicts.

Hypothesis 1 was supported in two instances when the data for the various categories of subjects were analyzed. Table 1 (columns 2 and 9) indicates that for subjects in publicly owned organizations and personnel specialists, there was a positive correlation between the attitudes of managers and the recruiting policies of their organizations.

#### 2. Hypothesis 2

When the total sample was considered, the data did not support the hypothesis. The data in the first column of Table 1 indicate that no significant relationships between the attitudes of subjects and any of the independent variables were found. However, when the data pertaining to the various categories of subjects were analyzed, a number of significant relationships were found. When subjects were classified according to the type of organization in which they were located, the hypothesis was largely confirmed for education. It was found that for subjects in publicly owned organizations,

TABLE 1  
PEARSON  $r$  COEFFICIENTS BETWEEN ATTITUDES AND INDEPENDENT VARIABLES

Independent variable	Total sample ( <i>n</i> = 217)	Type of organization			Skill level of employee recruited			Nature of subject's job		
		Public ( <i>n</i> = 33)	Primary/ manufact ( <i>n</i> = 92)	Service ( <i>n</i> = 92)	Unskilled/ semi-skilled ( <i>n</i> = 51)	Skilled ( <i>n</i> = 40)	Clerical ( <i>n</i> = 30)	Prof/ tech ( <i>n</i> = 40)	Personnel specialist ( <i>n</i> = 55)	General manager ( <i>n</i> = 110)
Personal										
Age	.0108	.2622	.0594	-.1990*	.3128**	-.3208*	.1158	.0183	.1513	-.0396
Education	.0889	.3669*	.1905*	-.1450	.1710	-.3326*	.2998	.3134*	.1030	.0331
Service	-.0085	.0786	.0596	-.0918	-.0569	-.1178	.2335	.1536	.1783	-.1425
Policy	.0986	.3885**	.0109	.0223	-.0498	.0847	.2703	.1157	.2311*	.0349
Organizational										
Age	-.1286*	-.2646	-.1071	-.0582	-.2019	.2039	-.0442	-.3609**	-.0409	-.2376**
Size	-.0238	.1285	-.0168	-.1300	.0831	-.2193	-.0567	.1512	.0364	-.1750*

\* Significant at .05 level.

\*\* Significant at .01 level.



and for privately owned primary and manufacturing organizations (Table 1, columns 2 and 3), the higher the level of their formal education, the more favorable were their attitudes toward hiring ex-convicts. Only one significant relationship was found between attitudes and age. It was found that for subjects employed in service organizations, attitudes towards hiring ex-convicts were inversely related to their age (Table 1, column 4). This is the converse relationship to that predicted in Hypothesis 2. No significant correlations were found between subjects' attitudes and their length of service.

Classification of managers according to the predominant skill level of the employees recruited by their organizations over the preceding two years revealed other data pertinent to Hypothesis 2. This information is presented in Table 1, columns 5, 6, 7, and 8. The hypothesized relationship between manager attitudes and their age was confirmed for organizations hiring unskilled and semiskilled laborers. However, the converse relationship held for organizations employing skilled labor. In organizations employing mainly skilled labor the converse relationship to that hypothesized was also found between managers' attitudes and the level of their formal education. The hypothesized relationship between attitudes and education held only for managers in organizations recruiting professional and technical employees. Once again, there were no significant correlations between managers' attitudes and their length of service with their organizations.

### 3. Hypothesis 3

The data failed to support this hypothesis. In fact, whenever a significant relationship was found between managerial attitudes and either organizational age or size, it was in the converse direction to that hypothesized. It was found that, for the sample as a whole, the attitudes of managers toward hiring ex-convicts were more negative the longer the organization had existed (Table 1, column 1). This relationship also held for managers in organizations recruiting mainly professional and technical personnel and for general managers (Table 1, columns 8 and 10). Only one significant correlation between organizational size and managerial attitudes was found. For organizations in which the general managers were responsible for the recruiting function, the attitudes of managers were negatively related to organization size (Table 1, column 10). This is, of course, the converse of the hypothesized relationship.

### D. CONCLUSION

Perhaps the most important finding of this study, from the point of view of persons concerned with increasing the effectiveness of efforts to increase



the rate of employment of ex-convicts, is that organizational policies with respect to the hiring of ex-convicts are independent of the attitudes of managers. Other empirical work conducted by the author suggests that other factors, such as the visibility of the organizations<sup>2</sup> and the extent to which they are subject to political pressure or control, are more likely to determine organizational policies towards hiring ex-convicts (3).

The implications of the finding that the attitudes of managers are inversely related to the age of their organizations were not immediately clear. Perhaps this finding is indicative of the managerial recruiting policies of organizations. Organizations may tend over time to become conservative and less prone to risk taking. This finding may also indicate that organizations tend increasingly to recruit and retain managers who subscribe to these conservative attitudes.

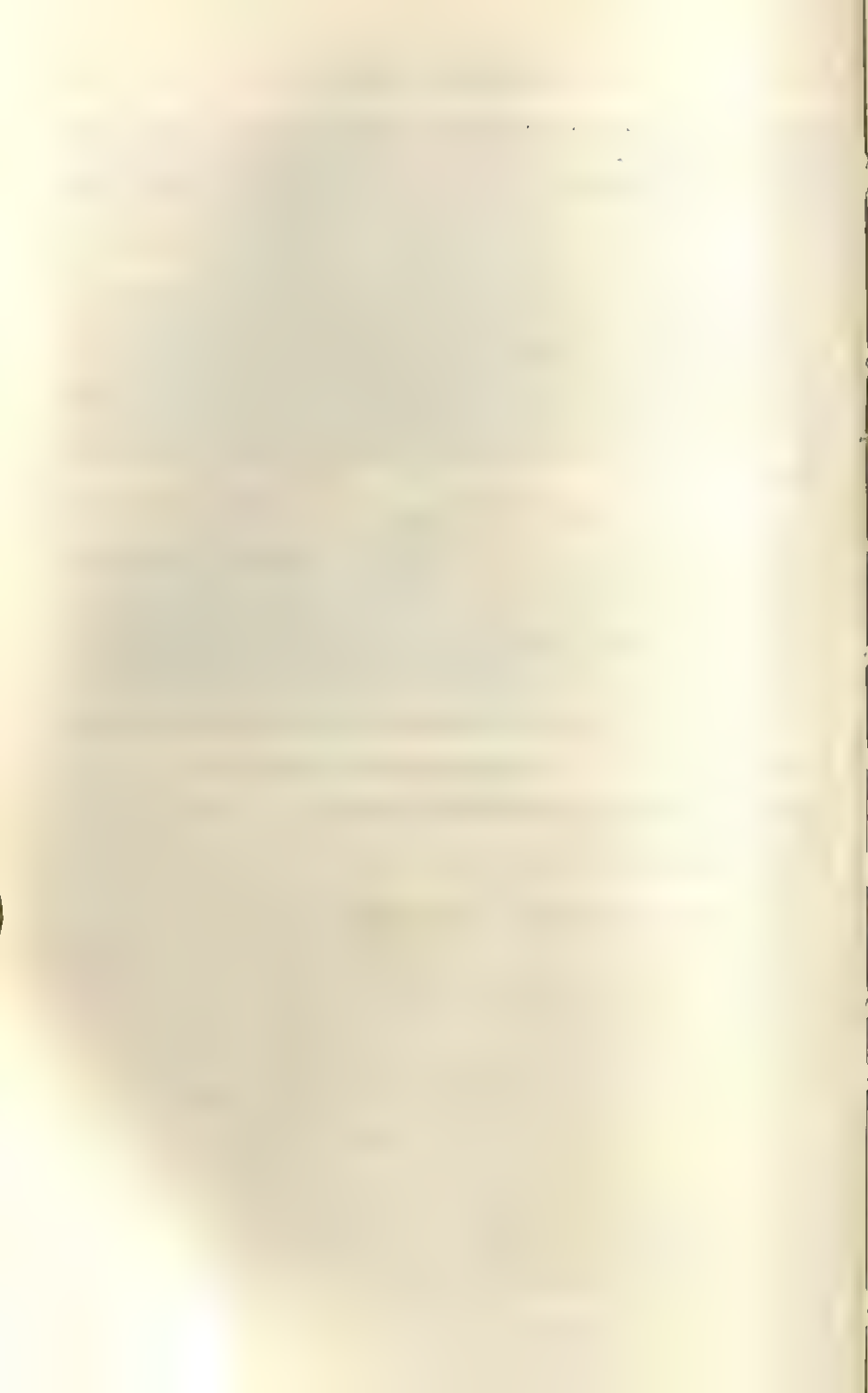
#### REFERENCES

1. KRETCH, D., CRUTCHFIELD, R. S., & BALLACHEY, E. L. *Individual in Society*. New York: McGraw-Hill, 1962.
2. MERTON, R. K. *Social Theory and Social Structure*. New York: Free Press, 1968.
3. NEDD, A. N. B. Some correlates of organization policies toward hiring ex-convicts. Unpublished manuscript, The University of Alberta, Edmonton, Canada, 1973.
4. NIE, N., BENT, D. H., & HULL, C. H. *SPSS, Statistical Package for the Social Sciences*. New York: McGraw-Hill, 1970.
5. RUBINGTON, E., & WEINBERG, M. S. *The Study of Social Problems*. New York: Oxford Univ. Press, 1971.
6. TRIANDIS, H. C. *Attitude and Attitude Change*. New York: Wiley, 1971.

*Faculty of Business Administration and Commerce*  
*The University of Alberta*  
*Edmonton, Canada T6G 2G1*

---

<sup>2</sup> The visibility of an organization is somewhat related to its size.



## EFFECTS OF INTERITEM ASSOCIATIVE STRENGTH, REHEARSAL, AND PROACTIVE INHIBITION ON THE RETENTION OF PAIRED-ASSOCIATE LISTS\*<sup>1</sup>

*University of Saskatchewan, Saskatoon, Canada*

---

JOHN A. MILLS

---

### SUMMARY

Eighty-six Ss were run in the proactive inhibition (PI) paradigm with the use of a  $2 \times 2 \times 2$  design with two levels of Interitem Associative Strength (IIAS), low and high; two of lists (one and two); and two of rehearsal (none and some). The Ss in the rehearsal condition were instructed to form associations between response terms and use them to retain stimulus-response connections. Recall was measured by a Modified Free Recall (MMFR) test followed by a Free Recall (FR) test; the retention interval was 24 hours. In the MMFR test, the hypothesis that high IIAS would cause decreased recall relative to low IIAS was confirmed, but the hypothesis that when IIAS was low, significant PI would be found only following rehearsal was not confirmed, so that previous findings by Houston were not replicated. In the FR test, the hypotheses that there would be significant PI, and that both rehearsal and high IIAS would benefit recall were confirmed. Hypotheses respecting the effectiveness of IIAS as a code were not confirmed.

---

### A. INTRODUCTION

A number of studies have shown that rehearsal enhances recall. For example, Rundus and Atkinson (12) found that recall of individual items in a free recall (FR) task increases as a positive function of the number of rehearsals per item. Johnson (6) showed that a period of rehearsal following 10 trials on a paired-associate (PA) task maintained a good level of performance on five criterial trials following the rehearsal and that the rehearsal conditions caused the gain of new associations. Postman and Stark (11)

---

\* Received in the Editorial Office on August 20, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The study was supported by grant no. APA 0279 from the National Research Council of Canada. The author is grateful for the assistance of Mrs. Dixie Weiss, Miss Mary-Alice Yunker, and Mr. Gordon Hemsley.

showed that telling Ss that a second list was to be presented following the first list in the retroactive inhibition (RI) paradigm resulted in a reduction in RI relative to conditions where Ss were not instructed. They attributed their finding to the fact that more Ss in the instructed groups rehearsed the first list than did Ss in the uninstructed groups.

On the other hand, Houston (4, 5) has shown that significant proactive inhibition (PI) occurs in PA learning only when Ss rehearse during a one-week retention interval, where rehearsal was defined in terms of either (a) explicit attempts to recall experimental material, or (b) an expectation that recall would be demanded, or (c) random thoughts about and associations with the experimental material. Houston explained his findings by saying that rehearsal in one-list conditions resulted in a strengthening of the experimental associations whereas, in the two-list conditions, rehearsal caused a gain in strength of the strong and a loss of strength of the weak items. The fact that rehearsal caused a lessening of interference in Postman and Stark's study and an increase in Houston's can be explained by saying that, in the former case, List 1 responses were rehearsed in the presence of appropriate cues and that the retention interval was much shorter than in Houston's study.

Weist (16) has evidence suggesting that the recoding necessary during rehearsal forms the basis for retrieval during recall. Associations between response terms could form the basis of a code in PA learning. Deese (3) has demonstrated that FR increases as a positive function of degree of interitem associative strength (IIAS) and Postman (9) has shown that long-term recall of serial lists of high IIAS is greater than those with zero IIAS. Both studies merely indicate that IIAS is potentially available as a code, since the gains of retention found can be attributed to the operation of a relatively passive associative mechanism. In this connection, it is noteworthy that Schwenn and Underwood (13) found that varying the strength of the association between pairs of response words did not affect rate of learning in PA lists, although they suspected that a negative effect was cancelling out a positive. However, using standard PA lists (which typically have low IIAS), Segal and Mandler (14) have shown that associations between "vertical" items beneficially affect FR when the standard (unidirectional) technique of PA learning is used and suggested that the division of items into left- and right-hand sets that is a consequence of the unidirectional method gives S effective retrieval cues during FR. Their work suggests that response-term IIAS could be used as an aid in recall during rehearsal, provided that the S was instructed to use it, since it should increase the strength of the "vertical"

associations and enhance the probability of retrieving stimulus items via the operation of backward associations.

The aim of the present study was to use a relatively long (24-hour) retention interval and to contrast the conditions used by Houston with conditions which might lead to a decrease in interference. A complicating factor is that, in most studies showing the beneficial effects of rehearsal on recall, FR tests have been used, whereas Houston used the test of modified free recall (MMFR) devised by Barnes and Underwood (1). Keppel (7) points out that MMFR tests are strongly affected by competition factors, so that the two tests give differing results. Following the technique of Ceraso and Henderson (2), the decision to use an MMFR test followed by an FR test was made. Although the MMFR test would give all Ss some opportunity to rehearse, it is known that PI persists after Trial 1 of relearning (10), so it was felt that some conditions would yield PI on the FR test. The predictions for the results of the MMFR test were as follows. In lists with low IIAS, it was predicted that PI would be found only in rehearsal conditions; in lists with high IIAS it was predicted that there would be both intra- and interlist interference, with the possibility that PI in such conditions would be higher than in conditions with low IIAS and rehearsal. The predictions for the results of the FR test were as follows. Although FR tests are not subject to specific competition effects, PI can be expected because a linguistic interfering task was used (7). It was predicted that high IIAS and rehearsal would lead to enhanced recall relative to low IIAS and no-rehearsal, that PI would be equally high in all non-rehearsal conditions, and that, whether IIAS was low or high, some reduction of PI would be seen in the low IIAS, rehearsal condition, and that a considerable reduction of PI would occur in the high IIAS, rehearsal condition.

## B. METHOD

### 1. *Subjects*

The Ss were 86 male and female university students aged 25 and under, obtained by advertising on the Saskatoon campus of the University of Saskatchewan. They were paid at a rate of \$2.00 per hour.

### 2. *Materials*

Nine-member PA lists were made up, conforming to the A-B, A-C paradigm. The stimuli were consonant trigrams with Witmer values of 50 percent (15, Appendix B). Two high IIAS lists were made up by selecting nine response



terms from Deese's (3) "Butterfly" and "Chair" lists. Low IIAS lists were made up by selecting nouns with low or zero associates from the Palermo-Jenkins (8) norms. In the two-list conditions, the first list was always low IIAS. All second lists were either high or low in IIAS. Warm-up lists were generated by randomly pairing patches of colored paper with the digits one to nine. The lists were mounted on Marietta memory drums with four different random orderings of the stimulus-response pairs.

### 3. Procedure

The design of the experiment was  $2 \times 2 \times 2$  factorial, with two levels of IIAS (low or high), two of lists (one or two), and two of rehearsal (none or some). The Ss were randomly assigned to one of the conditions. The Ss were told that the experiment involved the rote learning of the connection between a stimulus (a nonsense syllable) and a response (a common word). They were told that the stimulus would first appear alone followed by the stimulus paired with the response and that when the whole list had been presented once, their task was to say the correct response each time a stimulus appeared alone. The drums were set at a two:two-second rate with an eight-second intertrial interval. There were four trials of warm-up. First lists were learned to a criterion of nine correct anticipations on the criterial trial, second lists to a criterion of four.

The Ss in the nonrehearsal conditions were told that the experiment had been merely the preliminary to a more important experiment and they were asked to return 24 hours later. The Ss in the rehearsal conditions were told that they would have to recall the list(s) they had learned in 24 hours and were asked to rehearse. They were told to attempt to use associations between response terms during rehearsal. They were told that forgotten stimulus-response pairs could possibly be retrieved by the combined use of inter-response and backward associations.

The recall session consisted of an MMFR test, followed by an FR test in which the response had to be correctly assigned to lists by writing the response on a blank sheet of paper.

### C. RESULTS

There were no differences between pairs of lists within conditions (all  $F$  ratios  $< 1$ ), so the scores were pooled. Table 1 shows trials to criterion of the second lists for all conditions. An analysis of variance showed that Ss took significantly longer to reach criterion on high IIAS lists than they did on low ( $F = 11.50$ ,  $df = 1, 82$ ,  $p < .01$ ); there were no other effects. In order

TABLE 1  
MEAN TRIALS TO CRITERION OF SECOND LISTS AND MEAN SCORES  
ON THE MMFR AND FR TESTS

Rehearsal	Lists	IIAS			
		Low	N	High	N
Mean trials to criterion of second lists					
None	One	4.78(±3.96)	9	6.01(±3.28)	10
	Two	3.88(±2.17)	8	6.33(±4.24)	9
Some	One	3.00(±1.15)	13	7.10(±6.49)	10
	Two	3.69(±1.75)	13	4.86(±2.93)	14
Mean scores on the MMFR test (upper line) and FR test (lower line)					
None	One	4.67(±1.50) 5.44(±1.51)	9	4.90(±1.66) 7.20(±1.48)	10
	Two	3.00(±1.07) 4.25(±1.04)	8	3.78(±1.99) 6.33(±.87)	14
Some	One	6.23(±1.42) 6.92(±1.26)	13	4.40(±1.58) 6.90(±1.29)	10
	Two	4.38(±1.76) 5.62(±1.66)	13	3.64(±1.60) 6.50(±1.40)	9

Note: MMFR = modified free recall; FR = free recall; and IIAS = interitem associative strength.

to see if extent of recall and degree of learning were correlated, Pearson  $r$ s were run on total correct anticipations on the learning of second lists and scores on the MMFR and FR tests, pooling scores across rehearsal conditions within each level of the IIAS and list conditions. The only significant  $r$  occurred in the high IIAS, one-list condition ( $r = .54$ ,  $p < .01$ ). It was therefore concluded that it was unnecessary to adjust raw recall scores.

Rehearsal was defined as any expectation the  $S$  might have of being retested, any form of thinking about the material, and any form of explicit rehearsal, even of one association. Since  $S$ s did not necessarily obey the instructions given at the end of learning, reassignment to conditions was frequently necessary. There was therefore a possibility that there would be differences in ability to learn between the rehearsal and nonrehearsal groups, with the result that differences in extent of recall associated with differing amounts of rehearsal would have to be attributed to differing amounts of associative strength at the end of learning. Mean trials to criterion of all rehearsal groups was 4.50, while the score for nonrehearsal groups was 5.61. The difference between the means was not statistically significant ( $t = 1.45$ ,  $df = 84$ ).

Table 1 also shows the mean scores on the MMFR and FR tests. An analysis of variance showed that there was a significant effect for lists ( $F = 14.72$ ,  $df = 1, 78$ ,  $p < .001$ ). The overall mean for the one- and two-list conditions was 5.14 and 3.77, respectively; the amount of relative PI was 26.6 percent. There was a significant  $\text{IIAS} \times \text{Rehearsal}$  interaction ( $F = 6.50$ ,  $df = 1, 78$ ,  $p < .01$ ). The interaction is plotted in Figure 1. A Newman-Keuls test showed that when IIAS was low, the mean of the rehearsal conditions was significantly higher than that of the nonrehearsal ( $p < .05$ ), whereas the rehearsal and nonrehearsal means did not differ at high levels of IIAS. Furthermore, within rehearsal conditions, the mean of the high IIAS conditions was significantly lower than that for the low ( $p < .05$ ), whereas the means of the low and high IIAS conditions did not differ. In the non-significant  $\text{IIAS} \times \text{Lists}$  interaction, a Newman-Keuls test showed that significant PI occurred with low IIAS ( $p < .01$ ), but not with high IIAS. There was no significant difference between means for low and high IIAS within level of lists.

An analysis of variance of FR scores yielded a significant effect for IIAS ( $F = 15.97$ ,  $df = 1, 78$ ,  $p < .001$ ), for Lists ( $F = 10.27$ ,  $df = 1, 78$ ,  $p < .01$ ), and for Rehearsal ( $F = 5.31$ ,  $df = 1, 78$ ,  $p < .05$ ). The mean of the

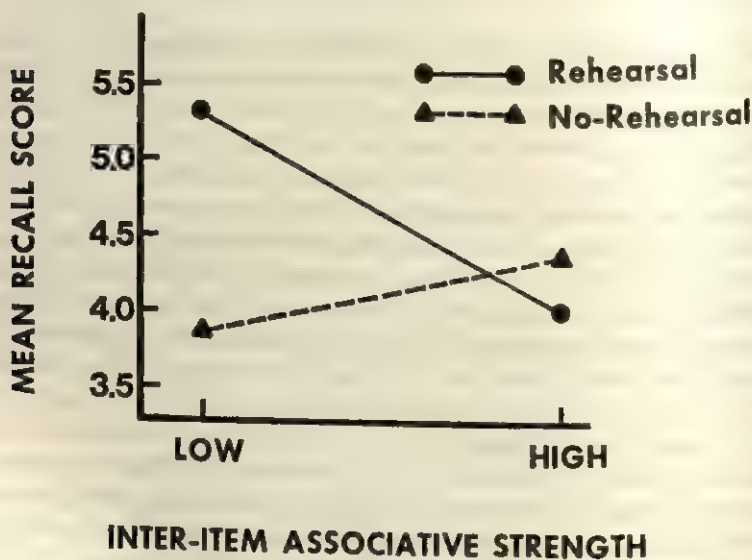


FIGURE 1  
INTERACTION BETWEEN IIAS AND REHEARSAL (MMFR TEST)

low IIAS conditions was 5.72, while that for the high was 6.72. The mean of the one-list conditions was 6.67 and of the two-list was 5.80; the amount of relative PI was 13 percent. The mean of the no-rehearsal conditions was 5.88, while that of the high was 6.46. There was also a significant IIAS  $\times$  Rehearsal interaction ( $F = 6.41$ ,  $df = 1, 78$ ,  $p < .01$ ). The interaction is plotted in Figure 2. A Newman-Keuls test showed that the low IIAS:no rehearsal mean was significantly lower than the others ( $p < .01$ ) and that the remaining means did not differ significantly.

#### D. DISCUSSION

With respect to MMFR, the hypothesis that the combination of high IIAS and rehearsal would lead to both intra- and interlist interference was sustained. The significant interaction between IIAS and rehearsal can be interpreted by saying that rehearsal raised recall in the low IIAS:rehearsal conditions, that degree of IIAS alone had no effect on recall, and that the effect of rehearsal in the high IIAS conditions was to depress recall because of intralist interference. The lack of significant PI when IIAS was high was probably due to intralist interference in one-list conditions, lessening the

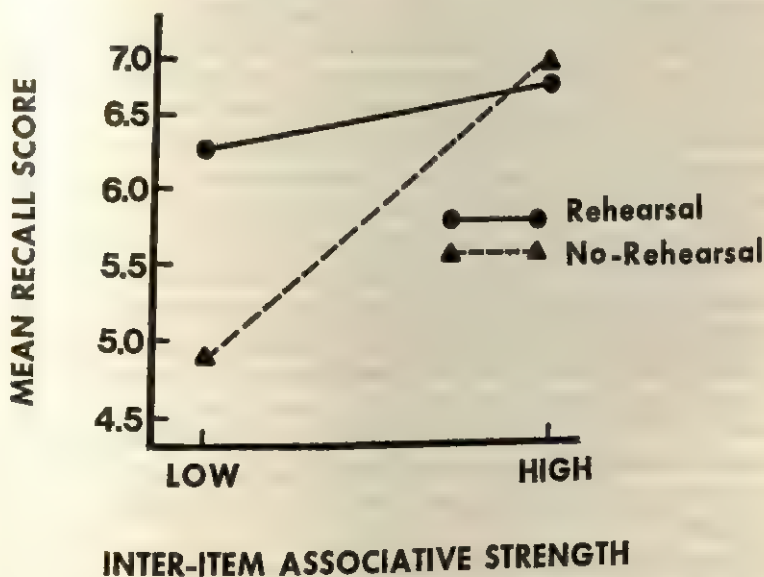


FIGURE 2  
INTERACTION BETWEEN IIAS AND REHEARSAL (FR TEST)

difference between them and two-list conditions. An examination of the means of the low IIAS conditions makes it possible to compare the results of the present experiment with those of Houston's. Both the failure to obtain a Lists  $\times$  Rehearsal interaction and the fact that relative PI was almost equally large in both the no-rehearsal and rehearsal conditions (35.8 and 29.7 percent respectively) shows that Houston's finding was not replicated. In his experiments, the effect of rehearsal relative to lack of rehearsal was to raise mean recall in the one-list conditions and lower it in the two-list, whereas in the present experiment the effect of rehearsal was to raise both means. A possible reason for the difference is that Houston used a higher criterion for second list learning than was used in the present experiment. It is therefore possible that his Ss had relatively large numbers of weak associations in the second list which were suppressed by competition from the first list during the retention interval. The Ss in the present experiment had to master relatively few associations during second list learning and it is possible that these were strong and resistant to interference and, at the same time, benefitted from the effects of rehearsal.

With respect to FR, the form of the IIAS  $\times$  Rehearsal interaction showed that either rehearsal or high IIAS led to a relative increase of recall but that high IIAS and rehearsal did not have a combined effect on recall, so that there was no evidence of a coding effect. There was thus a contrast with the findings respecting MMFR, where manipulation of IIAS alone had no effect and where high IIAS and rehearsal interacted to induce relatively poor recall. The beneficial effects of high IIAS on recall were presumably due to the lack of competition effects characteristic of FR tests, so that the results are comparable to those of Deese and of Postman, using single lists. The failure to obtain a significant IIAS  $\times$  Lists  $\times$  Rehearsal interaction showed, again, that the hypotheses concerning the effectiveness of IIAS as a code were not confirmed. The only indication that a coding process was at work came from comparing the amounts of relative PI generated in the various conditions, although the indication is weak, since the pattern of the results did not follow the predictions closely. High IIAS and rehearsal yielded a nonsignificant amount of relative PI (5.8 percent); high IIAS and no-rehearsal yielded an amount of relative PI that was barely significant (12.1 percent); low IIAS and rehearsal yielded 18.8 percent; and low IIAS and no-rehearsal yielded 21.9 percent.

In general, the results can be interpreted in terms of associative mechanisms. The fact that PI was greatly reduced in the high IIAS:rehearsal conditions could merely imply that rehearsal permitted the Ss to use their associative mechanisms relatively effectively, and it should be noted that high IIAS alone



led to a reasonable reduction of PI. The failure to obtain a coding effect can probably be attributed to the fact that Ss were unable to rehearse in the presence of appropriate cues. Although the maintenance of set alone has no effect on memory, as shown by such work as Johnson's (6), it would seem both from his work and from Postman and Stark's (11) that the presence of cues related to the experimental material (e.g., stimulus terms or the rotating blank memory drum) is important if rehearsal is to lead to relatively enhanced recall.

## REFERENCES

1. BARNES, J. M., & UNDERWOOD, B. J. "Fate" of first list associations in transfer theory. *J. Exper. Psychol.*, 1959, **58**, 97-105.
2. CERASO, J., & HENDERSON, A. Unavailability and associative loss in RI and PI. *J. Exper. Psychol.*, 1965, **70**, 300-303.
3. DEESE, J. Influence of inter-item associative strength upon immediate free recall. *Psychol. Rep.*, 1959, **5**, 305-312.
4. HOUSTON, J. P. Proactive inhibition and undetected retention interval rehearsal. *J. Exper. Psychol.*, 1969, **82**, 511-514.
5. ———. Proactive inhibition and undetected rehearsal: A replication. *J. Exper. Psychol.*, 1971, **90**, 156-157.
6. JOHNSON, R. E. Rehearsal effectiveness as a function of similarity of rehearsal activities to a paired-associate task. *J. Verb. Learn. & Verb. Behav.*, 1968, **7**, 439-445.
7. KEPPEL, G. Retroactive and proactive inhibition. In T. R. Dixon & D. L. Horton (Eds.), *Verbal Behavior and General Behavior Theory*. Englewood-Cliffs, N.J.: Prentice-Hall, 1968. Pp. 172-213.
8. PALERMO, D. S., & JENKINS, J. J. Word Association Norms: Grade School Through College. Minneapolis: Univ. Minnesota Press, 1964.
9. POSTMAN, L. The effect of inter-item associative strength on the acquisition and retention of serial lists. *J. Verb. Learn. & Verb. Behav.*, 1967, **6**, 721-728.
10. POSTMAN, L., & RILEY, D. A. Degree of learning and inter-serial interference in retention: A review of the literature and an experimental analysis. *Univ. Calif. Publ. in Psychol.*, 1959, **8**, 271-396.
11. POSTMAN, L., & STARK, K. Retroactive inhibition as a function of set during the interpolated task. *J. Verb. Learn. & Verb. Behav.*, 1962, **1**, 304-311.
12. RUNDUS, D., & ATKINSON, R. C. Rehearsal processes in free recall: A procedure for direct observation. *J. Verb. Learn. & Verb. Behav.*, 1970, **9**, 99-105.
13. SCHWENY, E. A., & UNDERWOOD, B. J. The effect of formal and associative similarity on paired-associate and free-recall learning. *J. Verb. Learn. & Verb. Behav.*, 1968, **7**, 817-824.
14. SEGAL, M. A., & MANDLER, G. Directionality and organizational processes in paired-associate learning. *J. Exper. Psychol.*, 1967, **74**, 305-312.
15. UNDERWOOD, B. J., & SCHULTZ, R. W. Meaningfulness and Verbal Learning. New York: Lippincott, 1960.
16. WEIST, R. M. The role of rehearsal: Recopy or reconstruct. *J. Verb. Learn. & Verb. Behav.*, 1972, **11**, 440-450.

Department of Psychology  
University of Saskatchewan  
Saskatoon, Canada S7N 0W0



## AN EXAMINATION OF VERBAL IMITATIVE PERFORMANCE IN YOUNG CHILDREN\*<sup>1</sup>

*University of Georgia*

---

REX FOREHAND AND HAROLD L. GARDNER

---

### SUMMARY

The present study examined the effects of *CA*, *MA*, and *IQ* on verbal imitation of five-year-old children. Three dependent measures of imitation were used: total, mimical, and conceptual. The results indicated that a transition occurs during the fifth year of life as young five-year-olds emitted more mimical and less conceptual responses than older five-year-olds. Furthermore, *IQ* is an important factor, as high *IQ* subjects, except when their *CAs* were less than five years, emitted fewer total and mimical responses than low *IQ* subjects.

---

### A. INTRODUCTION

In recent years imitation has received increasing recognition as an important method by which children learn behavior (1). Although motor behavior has been primarily examined in experimental studies of imitation, Rickard and his associates (6, 7, 8) have developed a methodological approach to the study of verbal imitative behavior: a subject listens to a tape-recorded model say words and, in response, the subject emits a word. Number of imitations of a critical word class, animal nouns, is examined.

In earlier studies by Rickard and his associates an imitative response was any animal word emitted by the subject. Holt, Rickard, and Ellis (5) have recently proposed that the word modeling responses can be divided into two categories: mimical and conceptual. The first type of response occurs when the subject emits the same word as the model. The second type of response,

---

\* Received in the Editorial Office on August 29, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This study was supported by Grant MH 22403-01 from the National Institute of Mental Health. Appreciation is expressed to the following agencies for making the study possible: Georgia Retardation Center (Athens branch), Headstart, Learning Tree Academy, the Model Cities program, and the University of Georgia Child Developmental Laboratory. Reprints may be obtained from the first-named author, at the address shown at the end of this article.

conceptual imitation, occurs when the subject emits a different word from the model but one in the same response class.

With children, particularly retardates, several critical variables have been identified in verbal imitation. Gardner and Forehand (4) found that institutionalized retardates imitate more than noninstitutionalized retardates. Forehand and Calhoun (2) demonstrated that retardates who manifest primarily mimical imitation also incorporate or accept the model's words as their own significantly more than retardates who emit conceptual imitation. Both Holt, Rickard, and Ellis (5) and Forehand, Robbins, and Brady (3) found that *IQ* is important in verbal imitation. Holt *et al.* reported that low *IQ* retardates (Mean *IQ* = 56) emitted significantly more mimical imitations than high *IQ* retardates (Mean *IQ* = 67), whereas Forehand *et al.* found that at the five-year *MA* level retardates (Mean *IQ* = 61) demonstrated significantly more mimical and total imitations than normals (Mean *IQ* = 102). Although the latter study compared normals and retardates at *MA* levels 5, 6, 7, and 8, the most consistent difference between the two *IQ* groups was at the five-year level.

The results of the Forehand *et al.* (3) study suggest that significant developmental changes in verbal imitation occur at the five-year-old level. The purpose of the present study was to examine the three types of verbal imitation (mimical, conceptual, and total) in children who have *CAs* and *MA*s in the five-year range. In Experiment I a  $2 \times 2$  factorial design with two *CA* levels (5-1 and 5-10 years) and two *IQ* levels (80 and 115) was used to examine the effects of *CA* and *IQ* on verbal imitation. In Experiment II a  $2 \times 3$  factorial design with two *MA* levels (5-0 and 5-10 years) and three *IQ* levels (62, 97, and 109) was used to investigate the effects of *MA* and *IQ*.

## B. METHOD

### 1. Subjects

*a. Experiment I.* The subjects were 36 children from preschool programs. Sixteen were from programs for disadvantaged children, nine were from a daycare center which served upper-lower-, and middle-class families, and 11 were from a middle-class kindergarten. Nine children were assigned to each of the following four groups: high *IQ*—low *CA*, high *IQ*—high *CA*, low *IQ*—low *CA*, and low *IQ*—high *CA*. The *IQ*s were determined by the Peabody Picture Vocabulary Test. Table 1 presents the characteristics of the subjects.

*b. Experiment II.* The subjects were 48 children from preschool programs and from the Georgia Retardation Center. Four were from programs for disadvantaged children, 21 were from daycare centers serving upper-lower

TABLE 1  
MEAN *CA*, *MA*, AND *IQ* OF SUBJECTS

Experiment I				Experiment II			
<i>Ss</i>	<i>CA</i>	<i>MA</i>	<i>IQ</i>	<i>Ss</i>	<i>CA</i>	<i>MA</i>	<i>IQ</i>
High <i>IQ</i>				High <i>IQ</i>			
High <i>CA</i>	5-9	7-2	115	High <i>MA</i>	4-9	5-10	111
Low <i>CA</i>	5-0	6-4	114	Low <i>MA</i>	4-5	4-11	107
Low <i>IQ</i>				Middle <i>IQ</i>			
High <i>CA</i>	5-10	4-8	80	High <i>MA</i>	5-11	5-10	98
Low <i>CA</i>	5-1	3-10	80	Low <i>MA</i>	5-5	5-1	95
				Low <i>IQ</i>			
				High <i>MA</i>	9-5	5-9	63
				Low <i>MA</i>	8-7	5-1	61

and middle families, six were from a middle-class kindergarten, and 17 were from the Georgia Retardation Center. Children were assigned to one of six cells in a  $2 \times 3$  factorial design: high *MA*—low *IQ*, high *MA*—middle *IQ*, high *MA*—high *IQ*, low *MA*—low *IQ*, low *MA*—middle *IQ*, and low *MA*—high *IQ*. The *MA*s and *IQ*s were determined by the Peabody Picture Vocabulary Test for all children except those from the Retardation Center. The Stanford-Binet and WISC were used with the latter children. Table 1 presents the characteristics of the subjects.

## 2. Apparatus

The apparatus was a cassette tape recorder and a cassette tape. The stimulus tape was prepared with the voice of a 28-year-old female who was not known by the subjects. The taped list consisted of 50 words. The first 10 words contained no animal words, the critical response class. The next 10 words contained two animal words, the next 10 contained four, the next 10 contained six, and the last 10 contained eight.

## 3. Procedure

After the subject was brought to the experimental room and seated in a chair, the experimenter gave the following instructions: "I want you to listen to this tape player. When you hear a voice say a word, you are to say a word. Say just one word. Say the first word you think of." Immediately following the instructions, the experimenter activated the tape player and the 50 stimulus words were presented at five-second intervals. The experimenter recorded the subject's responses. At the conclusion of the tape, the experi-



menter assured the subject that he had "done a good job," thanked him, and answered any questions.

### C. RESULTS

The data consisted of total imitations, conceptual imitations, and mimical imitations of words in the critical response class: animal words. Total imitations were defined as the total number of animal responses emitted by the subject. Mimical imitations were defined as the emission by the subject of the same critical class word as the taped model emitted. Conceptual imitations were defined as the emission of a nonmimical animal noun by the subject.

#### 1. Experiment I

Each of the three types of verbal imitation was examined by a  $2 \times 2$  analysis of variance with two between-subject factors (*CA* level and *IQ* level). Table 2 presents the mean number of total, mimical, and conceptual imitations at each *CA* and *IQ* level. The analysis of total imitations resulted in a significant effect at the borderline level for *IQ* ( $F = 3.36$ ,  $df = 1, 32$ ,  $p < .09$ ), indicating that low *IQ* subjects emitted more total imitations than high *IQ* subjects. The analysis of mimical imitations resulted in a similar borderline effect for the *IQ* factor ( $F = 2.88$ ,  $df = 1, 32$ ,  $p < .10$ ), again indicating more imitations by low than high *IQ* individuals. The analysis also revealed a significant effect for *CA* ( $F = 3.92$ ,  $df = 1, 32$ ,  $p < .07$ ) as the low *CA* group displayed more mimical responses than the high *CA* group. Similarly, the analysis of conceptual imitations resulted in a main *CA* effect ( $F = 3.32$ ,  $df = 1, 32$ ,  $p < .09$ ): high *CA* subjects displayed more conceptual imitation than low *CA* subjects.

#### 2. Experiment II

Each of the three types of verbal imitation was examined by a  $3 \times 2$  analysis of variance with two between subject factors (*IQ* level and *MA*

TABLE 2  
MEAN NUMBER OF THREE TYPES OF IMITATION AT EACH *CA* AND *IQ* LEVEL

Imitation	Low <i>CA</i>		High <i>CA</i>	
	Low <i>IQ</i>	High <i>IQ</i>	Low <i>IQ</i>	High <i>IQ</i>
Total	19.2	13.5	14.0	11.3
Mimic	19.2	12.6	11.8	9.1
Conceptual	0.0	.9	2.2	2.2

TABLE 3  
MEAN NUMBER OF THREE TYPES OF IMITATION AT EACH *CA* AND *IQ* LEVEL

Imitations	Low <i>MA</i>			High <i>MA</i>		
	Low <i>IQ</i>	Middle <i>IQ</i>	High <i>IQ</i>	Low <i>IQ</i>	Middle <i>IQ</i>	High <i>IQ</i>
Total	16.1	11.2	11.6	18.1	8.1	16.4
Mimic	13.2	9.5	10.1	16.3	6.6	15.8
Conceptual	2.9	1.7	1.5	1.8	1.5	.6

level). Table 3 presents the mean number of total, mimical, and conceptual imitations at each *MA* and *IQ* level. The analysis of total imitations resulted in a significant effect for *IQ* ( $F = 5.48$ ,  $df = 2, 42$ ,  $p < .01$ ). Further examination of the effect by way of the Least Squares Difference method (9) indicated that subjects in the middle *IQ* group produced fewer imitations than subjects in the low *IQ* group ( $t = 2.70$ ,  $df = 30$ ,  $p < .02$ )<sup>2</sup> and in the high *IQ* group ( $t = 2.30$ ,  $df = 30$ ,  $p < .05$ ). The analysis of mimical imitations also revealed an effect for *IQ* ( $F = 2.86$ ,  $df = 2, 42$ ,  $p < .08$ ). An examination of the effect by way of the Least Squares Difference method revealed that the middle *IQ* group produced fewer mimical imitations than the low *IQ* group ( $t = 2.31$ ,  $df = 30$ ,  $p < .05$ ). The difference between the middle *IQ* group and the high *IQ* group approached significance ( $t = 1.62$ ,  $df = 30$ ,  $p = .11$ ). The analysis of conceptual imitations did not result in any significant effects.

#### D. DISCUSSION

The results of Experiment I suggest that a transition in verbal imitation occurs during the fifth year of life: young five-year-olds employ mimical imitation significantly more than older five-year-olds. In contrast, the latter group uses conceptual imitation significantly more than the former group. Forehand *et al.* (3) hypothesized that conceptual imitation requires a higher level of cognitive functioning than mimical imitation. The former involves formulating the modeled word as a member of a response class and, subsequently, emitting another word within the same class, whereas a mimical response involves merely repeating the modeled word. The present findings suggest that higher level cognitive functioning appears in the latter part of year five.

The finding in Experiment I that low *IQ* subjects emitted more total and mimical imitations than high *IQ* subjects is similar to previous findings (3, 5). The finding in Experiment II that high *IQ* subjects, as did retardates,

<sup>2</sup> All *t* probability values are based on two-tail tests.

emitted more total and mimical verbal imitations than middle *IQ* subjects has not been previously reported. Since the middle and high *IQ* groups were matched on *MA*, the high *IQ* group had a *CA* (4-7 yrs.) below five years, while the middle *IQ* group had a mean *CA* (5-8 yrs.) in the five-year range. It would appear that a *CA* of at least five years is required for a decrease in total and mimical imitations to occur with an increase in *IQ*.

## REFERENCES

1. BANDURA, A. *Principles of Behavior Modification*. New York: Holt, Rinehart & Winston, 1969.
2. FOREHAND, R., & CALHOUN, K. Verbal imitation in retardates: Follow-up. *Percept. & Motor Skills*, 1973, **36**, 74.
3. FOREHAND, R., ROBBINS, B., & BRADY, C. P. Effects of *IQ* and mental age on verbal imitative performance of children. *J. of Psychol.*, 1973, **84**, 353-358.
4. GARDNER, H. L., & FOREHAND, R. The effects of institutionalization on word responses of retardates. Unpublished manuscript, University of Georgia, Athens, Georgia, 1972.
5. HOLT, M. M., RICKARD, H. C., & ELLIS, N. R. A note on word modeling in retarded adolescents. *Amer. J. Ment. Defic.*, 1972, **77**, 237-239.
6. RICKARD, H. C., & JOUBERT, C. Subject-model sexual status and observer performance. *Psychon. Sci.*, 1968, **10**, 407-408.
7. RICKARD, H. C., HOLT, M., & LATTAL, K. A. Reliability of word-modeling performance. *Psychol. Rep.*, 1970, **26**, 550.
8. RICKARD, H. C., ELLIS, N. R., BARNHART, S., & HOLT, M. Subject-model status and verbal imitative performance in kindergarten children. *Devel. Psychol.*, 1970, **3**, 405.
9. WINER, B. J. *Statistical Principles in Experimental Design*. New York: McGraw-Hill, 1962.

*Psychology Clinic*  
*Department of Psychology*  
*The University of Georgia*  
*Athens, Georgia 30601*

## FRUSTRATION RESPONSE CATEGORIES AND LEVEL OF HOSTILE EXPRESSION\*

*University of Missouri-Kansas City*

---

MORTON GOLDMAN AND CARYLN S. RHODS

---

### SUMMARY

Subjects were categorized on the basis of a Situational Inventory as having one of four response tendencies to annoying experiences: immediate action, deliberating action, accepting nonaction, and anxious nonaction. In the laboratory, subjects were placed in a frustrating situation and influenced to channel their responses in one of these same four modes. Measures were taken of hostility expression. Two findings were significant: (a) The level of expressed hostility, regardless of one's previous habitual preferred response, was affected by the way in which one dealt with annoyance in the current situation. (b) Habitual modes of responding to annoying situations were related to the amount of hostility that was expressed when faced with a frustrating agent. In both habitual and laboratory conditions, immediate action responses resulted in the highest level of hostile expression, and deliberating action responses resulted in the lowest level of hostile expression.

---

### A. INTRODUCTION

When faced with frustration are there habitual styles of responding that individuals will adopt? Would habitual ways of responding be related to the level of residual hostility felt toward a frustrating agent? If a given individual could be channelled into adopting a particular response style when frustrated, would his level of hostile feeling be affected by this mode of responding? These were the questions of concern in the present study.

The early work carried out to classify types of reactions to frustration was by Rosenzweig (7). However, his categories (extrapunitive, intrapunitive, and impunitive) were used generically rather than operationally. Little has been done in more recent times to categorize responses to frustration or to study the effects of a particular response style.

Factors that affect responses to frustration have been suggested by several

---

\* Received in the Editorial Office on August 30, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

investigators. The classical Frustration-Aggression Hypothesis by Dollard *et al.* postulated that the greater the strength of the goal-response sequence that is blocked, the greater the extent of interference, and the greater the number of frustrated sequences in a given period of time, the greater will be the tendency toward hostile expression (2). Berkowitz indicated that appropriate situational aggressive cues increased the probability that an overt aggressive response would occur when an individual was faced with frustration (1). Pastore reported that frustration that occurred because of nonspecific situational factors rather than from personally directed arbitrary factors produced less hostile expression (6). Psychoanalytic theory has suggested that expressing hostility in the face of frustration would reduce aggressive tendencies (catharsis), although more recent work has called this into question (4, 5).

The concerns of the present study were to see (a) if responses to frustration could be placed in a small number of categories; (b) given that categorization was possible, if people who have habitual response tendencies fitting into the prescribed categories would respond to frustration with differential levels of aggression; and (c) if individuals, regardless of typical ways of responding, when channelled into making categorical responses to frustration, would have residual levels of hostile feeling depending on the adopted mode of responding.

The current research began with the attempt to obtain some information about the types of situations that typically annoy college students and their characteristic way of responding to these situations. A questionnaire was administered to approximately 100 students enrolled in General Psychology, consisting of such questions as "List several situations which frustrated you in the past week or two," "How did you respond?" and "Were these typical ways in which you handle most problems?"

The responses to the preliminary questionnaire were analyzed, and an attempt was made to see if any systematic patterns appeared in the responses. On the basis of this analysis, four categories of habitual response tendencies were formulated, as follows:

(a) Immediate Action (I-A). The individual undertook some action at the first opportunity. The action was appropriate and had a realistic chance of success.

(b) Deliberating Action (D-A). The individual was more deliberate and contemplative and wanted to consider all possibilities before choosing a course of action, in order to give added assurance of success as well as to choose the most appropriate plan for all concerned.



(c) *Accepting Nonaction (Ac-NA)*. The person minimized the importance of the frustrating situation and tended not to get overly upset over it. He accepted the situation and hoped it would resolve itself with little involvement on his part.

(d) *Anxious Nonaction (Anx-NA)*. The person was definitely concerned about the situation but felt powerless to act.. He did not plan any constructive action to alleviate the situation but continued to be anxious about it.

The next stage of the study consisted of devising a Situational Inventory, using the situations most frequently cited on the questionnaire as arousing annoyance. A seven-item inventory was composed. Each item presented a situation which was annoying to students, and four choices were offered as possible solutions. Each of the choices represented one of the four response-tendency classifications (I-A, D-A, Ac-NA, and Anx-NA). The instructions for the Situational Inventory asked the student to imagine himself in each of the situations and to choose one of the four responses to indicate which he thought he would actually do.

The frustrating situations in the inventory consisted of disagreements with a roommate, a poor grade from a laboratory instructor, a teacher who was not dismissing class on time, a difficult final examination schedule, mis-managed finances, and the like.

According to their responses to the items on the Situational Inventory, subjects were categorized as having one of the four response tendencies. Two questions were posed for the current study: (a) Would a person's level of expressed hostility when faced with a frustrating agent be related to his habitual response tendency category? (b) Would a person's level of hostility when faced with a frustrating agent be affected by channelling him to adopt a particular response tendency regardless of his habitual response tendency?

## B. METHOD

### 1. *Experimental Subjects*

Subjects for this study consisted of 156 students enrolled in General Psychology who were administered the seven-item Situational Inventory. According to the responses chosen on the Inventory, they were categorized into one of the four response tendencies. A minimum of three responses of one type was required to categorize a person as having one of the four response tendencies. Most of the subjects had four or more responses of one type. Of the 156 subjects, 39 were categorized as I-A, 41 as D-A, 44 as Ac-NA, and 39 as Anx-NA.

## 2. *Experimental Design and the Assigning of Subjects to the Treatments*

A  $4 \times 4$  experimental design was employed. One main variable was the subjects' habitual response tendency as determined by the Situational Inventory. The four levels were I-A, D-A, Ac-NA, and Anx-NA. Subjects were categorized into these four levels as explained above. The other main variable was the laboratory manipulation which pressured the subjects to make given responses to frustration. A laboratory procedure was devised which first aroused hostile feeling through contact with a frustrating agent. The attempt was then made to have the subjects adopt one of four prescribed ways of handling their hostile feelings which duplicated the four levels of the response tendency categories. These four laboratory channelled responses provided the four levels for the second main variable.

Of the original sample of 156 subjects, approximately one-fourth fell into each response tendency category on the basis of the Situational Inventory. Within each of these response categories, approximately 10 subjects were assigned to each of the four laboratory treatments. That is, for those subjects who were classified into the I-A response tendency category, approximately 10 received the I-A laboratory treatment, 10 received the D-A laboratory treatment, 10 received the Ac-NA laboratory treatment, and 10 received the Anx-NA treatment. A similar distribution occurred for the three other response tendency categories (D-A, Ac-NA, and Anx-NA).

The laboratory sessions were run with four to six subjects at a time approximately eight weeks after the administration of the Situational Inventory, the sessions being repeated until all subjects had taken part in the study. At any one session, subjects from the different response tendency groups were present, as they were allowed to sign up for the laboratory sessions at their convenience. In a given session, any of the four laboratory treatments could be run as will be explained below.

## 3. *Laboratory Sessions and Experimental Treatments*

To participate in the laboratory session of this study the subjects entered the experimental room and were seated in assigned booths designed so that they could not see one another's desk area. They were told that the study was being carried out to formulate methods of teacher evaluation and to help evaluate student teachers, who would be instructing them in the making of objects by folding paper. As motivation, the subjects were informed that this task was closely related to the spatial ability sections used in standard *IQ* tests and gave indications of their ability to reason in abstract forms, and that recent research indicated that abstract reasoning ability was closely related to successful completion of college.

An experimental confederate, who was referred to as Mr. Johnson throughout the experimental sessions, acted as a student teacher. He was invited into the laboratory room to begin the paper-folding instructions. His dress was standardized: He always wore a black suit, white shirt, and tie. The instructions were for making four objects by folding square pieces of paper into a sailboat, a dustpan, a pinwheel, and a pyramid. The actual instructions<sup>1</sup> were given while the subjects attempted to make the object. The instructions were given at a predetermined rapid rate. To arouse the subjects' annoyance and hostility, preplanned mistakes and rapid speed of the instructions were used by the student teacher.

Upon completion of each object, subjects were requested to write a comment on a note form. They were told that in the normal classroom situation they would have many opportunities to talk to fellow students about the teacher's effectiveness as well as their feelings toward the teacher. Since the experiment was limited to one period, the natural classroom condition would be simulated by having them periodically pass notes to one another expressing opinions as to how well they thought the teacher was doing.

The notes written by the subjects were collected, and under the guise of shuffling the notes so that subjects supposedly would receive a note from a different subject each time, the experimenter substituted previously prepared standardized notes. The content of the prepared notes permitted the introduction of the four experimental treatments.

*a. Immediate Action (I-A) treatment.* The notes that were returned to the subjects in this treatment stated that the subjects "wanted definitely to do something" about the poor teaching and that they would be getting together at the end of the session to talk to him about what he was doing wrong.

*b. Deliberating Action (D-A) treatment.* The notes returned to the subjects in this treatment stated that the subjects wanted to "plan to do something" about the poor teaching and that they would be getting together at the end of the session to formulate some plan.

*c. Accepting Nonaction (Ac-NA) treatment.* The notes returned to the subjects in this treatment stated that it was "not worth getting annoyed about" the poor teaching, that it was not worth becoming too involved or getting too excited about, and things would work out.

*d. Anxious Nonaction (Anx-NA) treatment.* The notes returned to the subjects in this treatment stated that the subjects "did not know what to do"

---

<sup>1</sup> Instructions, timing, and intentional errors used in the present study are the same as those used in a prior study (3).

about the poor teaching, were unhappy and worried about the situation, but felt there was nothing they could do.

#### 4. *Measurement of the Dependent Variable*

After the instructions for the fourth object had been completed and the final prepared notes distributed and read by the subjects, the student teacher was dismissed from the experimental room, and subjects were asked to complete a Teacher Rating Form. This form consisted of five-point rating scales (1-poor, 2-below average, 3-about average, 4-above average, and 5-good) dealing with teacher presentation, appearance, timing of instructions, and preparation. The subjects were also asked to give the teacher a letter grade of A, B, C, D, or F. These four scales plus the assigned letter grade were combined to constitute the teacher evaluation rating.

A five-point scale was also used for items asking the subjects the extent to which they wanted this person as a regular teacher, enjoyed or were frustrated working under the student teacher, and wanted to reward or punish the student teacher.

### C. RESULTS

This study consisted of initially giving a Situational Inventory, and according to the way in which the subjects responded to the items, they were separated into four response tendency categories. In the laboratory the subjects were frustrated by poor instruction from a student teacher. By a laboratory procedure their reactions were channelled into one of four different patterns which paralleled the response categories of the Situational Inventory.

It is possible that, because of the superior ability of some subjects, the teacher's instructions might have been adequate for them to learn the material and perform at an effective level; hence, frustration might not have occurred for these subjects. The paper objects on which the subjects received instructions had successive stages of completion, and this afforded a means by which to assess their performance. It was arbitrarily decided that students who received 12 or more points out of a total possible of 14 should be eliminated from the main analyses because they were not sufficiently affected by the frustrating aspects of the situation. By use of this criterion, 10 subjects from the I-A laboratory treatment, 9 from the D-A, 14 from the Ac-NA, and 9 from the Anx-NA were eliminated from the data analyses.<sup>2</sup>

<sup>2</sup> The eliminated subjects checked average or above for all the items on the Teacher Rating Form; that is, they evaluated the student teacher somewhat positively, were not frustrated, and did not wish to punish him.



To begin the analyses, a mean score for each subject was calculated on the basis of the five teacher evaluation items completed. The "Evaluation" row of the top half of Table 1 presents the means of the teacher evaluation scores for the subjects categorized by the Situational Inventory into the four response tendencies, ignoring the four laboratory treatments. The "Evaluation" row of the bottom half of the table presents the mean scores for the subjects in the four laboratory treatments, ignoring the four response tendencies. A  $4 \times 4$  analysis of variance was run, yielding an  $F = 2.78$  ( $df$  of  $3/98$ ,  $p < .05$ ) for the response tendency factor and an  $F = 2.70$  ( $df$  of  $3/98$ ,  $p < .05$ ) for the laboratory treatment factor. The interaction was nonsignificant, the  $F$  being less than one.

The "Teacher again" rows of Table 1 show the means obtained for the item asking the subjects whether they would want the student teacher as a regular teacher, for the four response tendencies and four laboratory treatments, respectively. For the  $4 \times 4$  analysis of variance test, an  $F = 3.08$  ( $df$  of  $3/98$ ,  $p < .05$ ) was obtained for the laboratory treatment factor. The interaction and response tendency factor yielded nonsignificant  $F$ s.

The "Frustrated" rows of Table 1 similarly present the means for the items asking the subjects the extent to which they felt frustrated working

TABLE 1  
TEACHER RATING MEANS AND RANKINGS FOR LABORATORY TREATMENT  
AND RESPONSE TENDENCIES

Teacher ratings	Response categories							
	I-A		D-A		Ac-NA		Anx-NA	
	Mean	Rank	Mean	Rank	Mean	Rank	Mean	Rank
<i>Situational Inventory response tendencies</i>								
	(n = 29)		(n = 32)		(n = 31)		(n = 22)	
Evaluation	13.95	1	15.80	4	15.74	3	15.71	2
Teacher again	1.95	1	2.38	3	2.35	2	2.44	4
Frustrated	2.02	1	2.21	4	2.09	2	2.11	3
Punish teacher	2.48	1	2.75	3	2.76	4	2.72	2
Average rank		1		3.5		2.8		2.8
<i>Laboratory treatments</i>								
	(n = 27)		(n = 29)		(n = 28)		(n = 30)	
Evaluation	15.19	3	16.39	4	15.18	2	14.44	1
Teacher again	2.06	1	2.65	4	2.26	3	2.14	2
Frustrated	1.87	1	2.33	4	2.06	2	2.18	3
Punish teacher	2.44	1	2.91	4	2.74	3	2.63	2
Average rank		1.5		4.0		2.5		2.0

Note: Low score indicates a poor rating. I-A = Immediate Action; D-A = Deliberating Action; Ac-NA = Accepting Nonaction; Anx-NA = Anxious Nonaction.



under the student teacher. For the  $4 \times 4$  analysis of variance test, an  $F = 2.87$  ( $df$  of  $3/98$ ,  $p < .05$ ) was obtained for the laboratory treatment factor. The interaction and response tendency factor were nonsignificant.

Finally the "Punish teacher" rows of Table 1 give the means for the item asking the subjects the extent they wanted to punish the teacher. On the  $4 \times 4$  analysis of variance test an  $F = 2.77$  ( $df$  of  $3/98$ ,  $p < .05$ ) was obtained for the laboratory treatment factor. The interaction and response tendency factors again were nonsignificant.

To summarize, the means obtained on the teacher evaluation items, the item dealing with whether the subject would want the student teacher as a regular teacher, and the items dealing with the extent to which the subjects felt frustrated and wanted to punish the teacher, were all significantly different for the four laboratory treatments.

For the response tendency factor, significant results were found on the teacher evaluation items, but not for the other items.

Table 1 also presents the rank order of the means for each of the above ratings for the four response tendencies and for the four laboratory treatments. A fairly consistent pattern occurred in the rank order of the means for the four response tendencies, with the I-A treatment giving the lowest ratings and the D-A giving the highest. A Kendall coefficient of concordance run for the ranks was found to be significant at the .05 level.

This same pattern was obtained for the rank order of the means for the laboratory treatments, again with the I-A treatment giving the lowest ratings and the D-A giving the highest. A Kendall coefficient of concordance run for these rankings was also significant at the .05 level.

Table 1 can also be examined to determine how the subjects felt toward the student teacher regardless of laboratory treatment or initial response tendency. Most subjects on the teacher evaluation items rated the teacher below average, the mean letter grade assigned being between C and D; and subjects indicated that they would not want this student teacher as a regular teacher. Also, the subjects indicated that they were frustrated by the student teacher, and felt some aggression toward him; i.e., they wanted to punish him because the instructions kept them from producing a good performance.

#### D. DISCUSSION

One of the major results of the current study was that the subjects' level of expressed hostility was affected by the particular category into which they were pressured to channel their reactions to annoyance in the laboratory treatments, regardless of their typical way of responding. The highest level

of hostile feelings occurred in the I-A treatment, and the lowest level occurred in the D-A treatment. Examination of the experimental manipulations which produced this difference shows that the prepared notes indicate as follows: (a) in the I-A treatment an attempt would be made immediately following the laboratory session to correct the teacher; and (b) in the D-A treatment, subjects would be getting together to plan a possible action (which could turn out to be no action). One explanation for the higher level of expressed hostility in the I-A treatment is that the subjects were generating an annoyance set preparing them for the immediate confrontation with the teacher. It is as though one were anticipating a possible struggle and building up the necessary strength of conviction to justify the action to be taken. In the D-A treatment, the notes did not suggest an immediate confrontation with the teacher but rather a meeting with the other subjects for a discussion about some possible future action. Being concerned with discussion and decision making in a group may serve to hold back or dissipate hostile feelings, at least until further intentions can be clarified. The intent in this case may have been to solve the problem rationally with deliberation, rather than to have a possibly aggressive confrontation with the frustrating agent, as in the I-A condition.

In the Ac-NA treatment, the communications that subjects received stated they should not become overly concerned, that it was not worth getting excited about, and that things would work out. In spite of these "holding the lid on" notes, the students' ratings were lower in this treatment than in the D-A treatment. A possible explanation may be that in the Ac-NA treatment the notes were saying "Don't get annoyed," but there was no behavioral response (consequence) for either being annoyed or not being annoyed. In the D-A treatment the notes received by the subjects attempted to foster negative feeling toward the teacher but, as explained above, more involvement may have been intimated on the part of the students than was warranted. Negative feelings might be more readily expressed when no actions were to be planned as a consequence of this negative expression.

The expressed hostility level in the Anx-NA treatment also fell between that of the D-A and I-A treatments. This would seem to be in line with the above explanations. Subjects in the Anx-NA treatment did not receive notes telling them to keep down their negative feelings about the instructor. The tenor of the notes in the Anx-NA treatment was to be concerned and unhappy but provided no anticipation of corrective action. Here again, there was no expected confrontation with the student teacher as was the case for the I-A treatment subjects, and thus less negative feeling was generated.

The results also showed that subjects who were categorized by the Situational Inventory into the four response tendencies expressed different levels of hostility toward the teacher. Here also, the greatest difference of levels of expressed hostility occurred between the I-A and D-A treatments, with the highest level occurring in the I-A treatment and the lowest level occurring in the D-A treatment. Thus, the study found a consistency between the habitual modes of dealing with annoyances and the induced laboratory treatments. In both instances subjects who were categorized as having I-A response tendencies and subjects who were channelled by the laboratory treatment into handling their annoyance in an I-A fashion expressed the most negative feelings toward the teacher. Similarly, in both instances the D-A subjects expressed the lowest level of hostile feeling toward the teacher.

The results obtained with the use of the Situational Inventory categories implied a relationship (correlation) between habitual ways of responding and levels of hostile feeling. However, these results, combined with the data obtained with the laboratory manipulations, implied a causal relationship between frustration and level of hostile feeling.

For the most part, previous research has dealt with an individual responding to a frustrating situation without any accompanying group influence. The current study was unique in showing that a particular response tendency made by a group toward a frustrating agent can affect the level of hostile expression of an individual in that group.

#### REFERENCES

1. BERKOWITZ, L. The frustration-aggression hypothesis revisited. In L. Berkowitz (Ed.), *Roots of Aggression*. New York: Atherton Press, 1969. Pp. 1-28.
2. DOLLARD, J., DOBB, L., MILLER, N., MOWRER, O., & SEARS, R. *Frustration and Aggression*. New Haven, Conn.: Yale Univ. Press, 1939.
3. GOLDMAN, M. Alternative classroom standards concerning management of hostility and effects on student learning. Unpublished Doctoral dissertation, University of Illinois, Urbana, 1955.
4. GOLDMAN, M., KECK, J. W., & O'LEARY, C. J. Hostility reduction and performance. *Psychol. Rep.*, 1969, **25**, 503-512.
5. MAILLICK, S. K., & MCCANDLESS, B. R. A study of catharsis of aggression. *J. Personal. & Soc. Psychol.*, 1966, **4**, 591-596.
6. PASTOR, N. The role of arbitrariness in the frustration-aggression hypothesis. *J. Abn. & Soc. Psychol.*, 1952, **47**, 728-731.
7. ROSENZWEIG, S. A general outline of frustration. *Charac. & Personal.*, 1938, **7**, 151-160.

Department of Psychology  
University of Missouri-Kansas City  
Kansas City, Missouri 64110

## RISK TAKING BY INDIVIDUALS AND INFORMAL GROUPS WITH THE USE OF INDUSTRIAL PRODUCT PURCHASING SITUATIONS AS STIMULI\*

*University of Cincinnati*

---

PETER H. REINGEN

---

### SUMMARY

Individual *versus* group risk taking was examined by using a novel set of industrial product purchasing situations as stimuli. The set incorporated situations that varied along a dimension called "normative risk." Situations represented either low-normative risk, medium-normative risk, or high-normative risk. Normative risk was defined in a decision theoretical context. The set was administered to introductory business students at a large university in the United States. The results of the study showed that groups were not invariably riskier than individuals; rather the amount of risky shift following group discussion was negatively related to the level of normative risk. This overall conclusion held true also in relation to the degree of perceived riskiness of the situations, since a positive relationship between normative risk and perceived risk was established. The theoretical implications of the results are discussed.

---

### A. INTRODUCTION

Rarely in the history of social psychology has a single study stimulated as much research as the Master's thesis by Stoner (11), which reported the "risky shift." Its conclusion that groups take riskier positions after discussion than the average of individual decisions prior to group discussion was contrary to the body of opinion and research that suggested that groups were either more conservative than individuals or that groups produced an averaging effect (1, 13). Since then a large amount of research (7) has investigated the effects of group discussion on responses to questionnaire items on which subjects recommend the amounts of risks to be taken by hypothetical persons facing various life dilemmas (6, Appendix E).

Despite the large number of replications with a variety of populations,

---

\* Received in the Editorial Office on September 4, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.



evidence for the risky shift phenomenon is severely limited by the fact that the overwhelming majority of studies have used the same set of the Kogan and Wallach (6) hypothetical life-dilemma situations. When different hypothetical risk situations have been used, it has been found that group discussion can lead not only to a risky shift but also to a conservative shift or to no shift at all (4, 5, 12). This conclusion was tested in the present study with a novel set of stimuli: namely, industrial product purchasing situations which varied along a dimension called "normative risk." Normative risk was defined in a decision theoretical context. Consider a decision-making matrix with two alternative courses of action and two states of nature. One of the courses of action is more conservative, in that its expected cost is not affected by the states of nature. The expected cost of the other course of action is affected by the states of nature. This course of action is riskier, but potentially more rewarding, if the more desirable of the two states of nature is the true state of nature. Normative risk refers to that lowest level of probability of the desired state of nature for which the expected cost of the conservative course of action exceeds the expected cost of the riskier alternative course of action.<sup>1</sup> Previous risk-taking instruments involving life-dilemma situations can be criticized on the grounds that they provide insufficient information to calculate the probabilities. Thus, they were considered normatively deficient (8).

The risk-taking instrument of the present study involved three normative-risk levels: low, medium, and high. Following the Kogan and Wallach (6) diffusion-of-responsibility explanation of the risky-shift phenomenon, one would predict a risky shift following group discussion, no matter what the normative levels of risk. Under the assumption that perceived risk by subjects approximates normatively prescribed risk taking, one might even predict that the higher the level of normative risk, the greater the opportunity for diffusion of responsibility, and the greater the risky shift following group discussion. However, there is evidence suggesting a negative relationship between perceived seriousness of consequences of life-dilemma situations and degree of shift toward risk (3). Therefore, one is led to quite different predictions. It was hypothesized that the amount of risky shift following group discussion would be negatively related to the level of normative risk. Consequently, and necessarily so, a positive relationship between perceived risk and normative risk was expected.

---

<sup>1</sup> It is assumed that utility and money are linearly related over the range of possible outcomes. See Richmond (9) for a more detailed discussion of decision making under risk.



## B. METHOD

1. *Experimental Condition*

a. *Subjects.* Participants were 61 unpaid introductory business students at the University of Cincinnati who received partial credit for their participation.

b. *Risk-taking instrument.* The risk-taking instrument consisted of three purchasing situations involving industrial goods. Each purchasing situation represented either low-normative risk, medium-normative risk, or high-normative risk. The central person in each situation had to choose between two courses of action, one of which was riskier than the other but also more rewarding if successful. For each situation the subject had to indicate the lowest probability of the more desirable state of nature he would accept before recommending that the potentially more rewarding alternative be chosen. The probabilities listed are 10, 20, 30, 40, 50, 60, 70, 80, and 90 chances of success in 100, plus an initial category (scored as zero) and a final category (scored as 100) in which the subject can accept or refuse to recommend the risky alternative no matter how its likelihood of success.

The purchasing situation involving low-normative risk (LNR) is given below as administered to the subjects.

Mr. Slocum is the purchasing manager of the Cutler Cordage Company. The Cutler Cordage Company is a large, well-established company manufacturing many types of Manila rope and cordage products. Abaca fiber is used in the manufacture of Manila rope and accounts for more than 85% of the dollars spent by the purchasing department. To meet the demand for Manila rope, Mr. Slocum needs a supply of 41,510 pounds of Abaca fiber. Mr. Slocum can order this amount of Abaca fiber now for \$.45 per pound. Because of the speculative nature of the market for Abaca fiber, however, the price for Abaca fiber could drop to \$.21 per pound or increase to \$.55 after two months. The company has enough inventory of Abaca fiber to wait two months. Inventory costs are negligible, and instant delivery is assured. Mr. Slocum can either order now or in two months.

In this LNR situation, the minimum normative probability level is 30 chances in 100.<sup>2</sup>

Similarly, the purchasing situation representing medium-normative risk (MNR) required a minimum normative probability of 50 chances in 100 for the expected cost of the conservative alternative to exceed the expected cost of the riskier alternative; the purchasing situation depicting high-normative risk (HNR) necessitated a minimum normative probability of 70 chances in 100. These situations are summarized below, respectively.

<sup>2</sup> Expected cost = .30 (\$.21) + .70 (\$.55) = \$.448 < \$.45.

MNR: A purchasing manager is facing a decision whether to buy one of two machines. The purchase of the more expensive machine could amount to a substantial discounted cost saving (e.g., less repairs) over its lifetime use.

HNR: A purchasing agent at a university is facing a decision to award a contract for the purchase of desks to one of two competitive bidders. The company representative of the initially more expensive bidder offers a reduction in bid price one day before the contract is to be awarded. Bid prices were suggested prices.

The worst possible outcome was held constant over all three situations. The instrument showed adequate test-retest reliability for a separate similar group of students used as control. The resulting product-moment correlation was .79 for the instrument as a whole.

c. *Procedure.* Ss were told to participate in an experiment on decision making. After receiving instructions from E, Ss individually completed the risk-taking instrument. They were allowed 10 minutes. After completion, E collected the questionnaires and distributed new copies of the risk-taking instrument. Ss were randomly assigned to groups.<sup>8</sup> They discussed and arrived at unanimous decisions for each item. Groups were allowed 15 minutes to reach consensus decisions.

## 2. Control Condition

Fifteen similar subjects served as controls. Control subjects completed the risk-taking instrument in two sessions one week apart. Instructions and time constraints were those of the experimental prediscussion condition.

## C. RESULTS

### 1. Normative Risk versus Perceived Risk Experimental Condition

Individual prediscussion risk-taking positions were significantly more conservative than normatively prescribed risk taking for the LNR situation ( $\bar{X} = 47.19$ ;  $t = 5.63$ ;  $p < .001$ ), and they were significantly riskier than the normative risk-taking level for the HNR situation ( $\bar{X} = 61.64$ ;  $t = -2.88$ ;  $p < .01$ ). The consensual group decisions after discussion were not significantly different from the normative risk-taking level for the LNR situation ( $\bar{X} = 37.37$ ;  $t = 1.86$ ;  $p < .10$ ), but for the HNR situation, groups were significantly more conservative than the normative risk-taking level ( $\bar{X} = 78.95$ ;  $t = 2.22$ ;  $p < .05$ ). Neither the individual prediscussion decisions ( $\bar{X} = 54.65$ ) nor the group decisions ( $\bar{X} = 49.40$ ) differed significantly from the normative risk-taking level for the MNR situation. Individuals

<sup>8</sup> There were 15 three-person groups and four four-person groups.

appeared to be less willing to take normatively prescribed extreme risk-taking positions than groups. Although several of the differences between normative risk and perceived risk were significant, the results did, however, confirm the expectation of a positive relationship between the perception of riskiness of a given situation and the normative risk level of that situation.

## 2. Individual versus Group Risk Taking

Table 1 examines the significance of the risk-taking differences between the mean of the prediscussion individual decisions made by the members of each group and that group's consensual decisions. Tests were carried out for each item separately and for the three items combined. Since larger scores indicate less risk taking, a negative difference indicates a shift in the risky direction. A *t* test for related measures was used to determine whether the scores were significantly different from zero.

For all three items combined, consensual decisions of the groups were less risky than the individual prediscussion positions of the members of each group. However, this shift in the conservative direction was not significant. When each item was considered separately, a significant risky shift occurred for the LNR situation. Groups shifted in a significant conservative direction for the HNR situation, while no significant shift materialized for the MNR situation. Therefore, the results were as expected. The amount of risky shift was negatively related to the level of normative risk and, thus, negatively related to perceived risk. The control subjects did not exhibit a significant shift in either the conservative or risky direction at the .05 level.

## D. DISCUSSION

The results of this study seem to indicate that groups can shift in a risky direction, conservative direction, or not shift at all, depending upon whether

TABLE 1  
SIGNIFICANCE OF RISK-TAKING DIFFERENCES BETWEEN MEAN OF PREDISCUSSION  
INDIVIDUAL DECISIONS FOR A GROUP'S MEMBERS AND  
GROUP'S CONSENSUAL DECISIONS

Item	Mean difference	df	s	t
All combined	1.12	18	11.00	.45
MNR level	— 5.17	18	15.67	—1.44
LNR level	— 9.82	18	19.19	—2.33*
HNR level	17.31	18	18.90	3.99**

Note: MNR = medium-normative risk; LNR = low-normative risk; HNR = high-normative risk.

\*  $p < .05$  (two tailed).

\*\*  $p < .001$  (two tailed).

a decision-making situation depicts low-normative risk, high-normative risk, or medium-normative risk, respectively. Since a positive relationship between normative risk and perceived risk was established by the data, the above overall conclusion apparently holds true in relation to the degree of perceived riskiness of the situations. Rather than speaking of a risky shift phenomenon, it is suggested that one refer to phenomena of shifts along a risk dimension.

If group discussion can produce risky shifts and conservative shifts, as well as no shifts, and if future experiments demonstrate that the outcome of risky shift experiments can be predicted as a function of normative and/or perceived risk levels, what theory would best account for these results? It is suggested here that at least two factors are involved. First, there might be a bandwagon effect. When group members discover that most of them are leaning in the same direction on an item, this confirmation creates affective pressure for perceiving that particular level of risk as appropriate. This bandwagon tendency can carry group decisions to more extreme choices. Second, group discussion can lead to cognitive change in that it leads to a better understanding of such things as the stimuli characteristics, the value of positive outcomes, and the costs of negative consequences.

The observed differences between normatively prescribed decision-making behavior and actual decision-making behavior seem to confirm previous findings. Subjects may "satisfice" rather than optimize choice (10), and they may not use all information available (2).

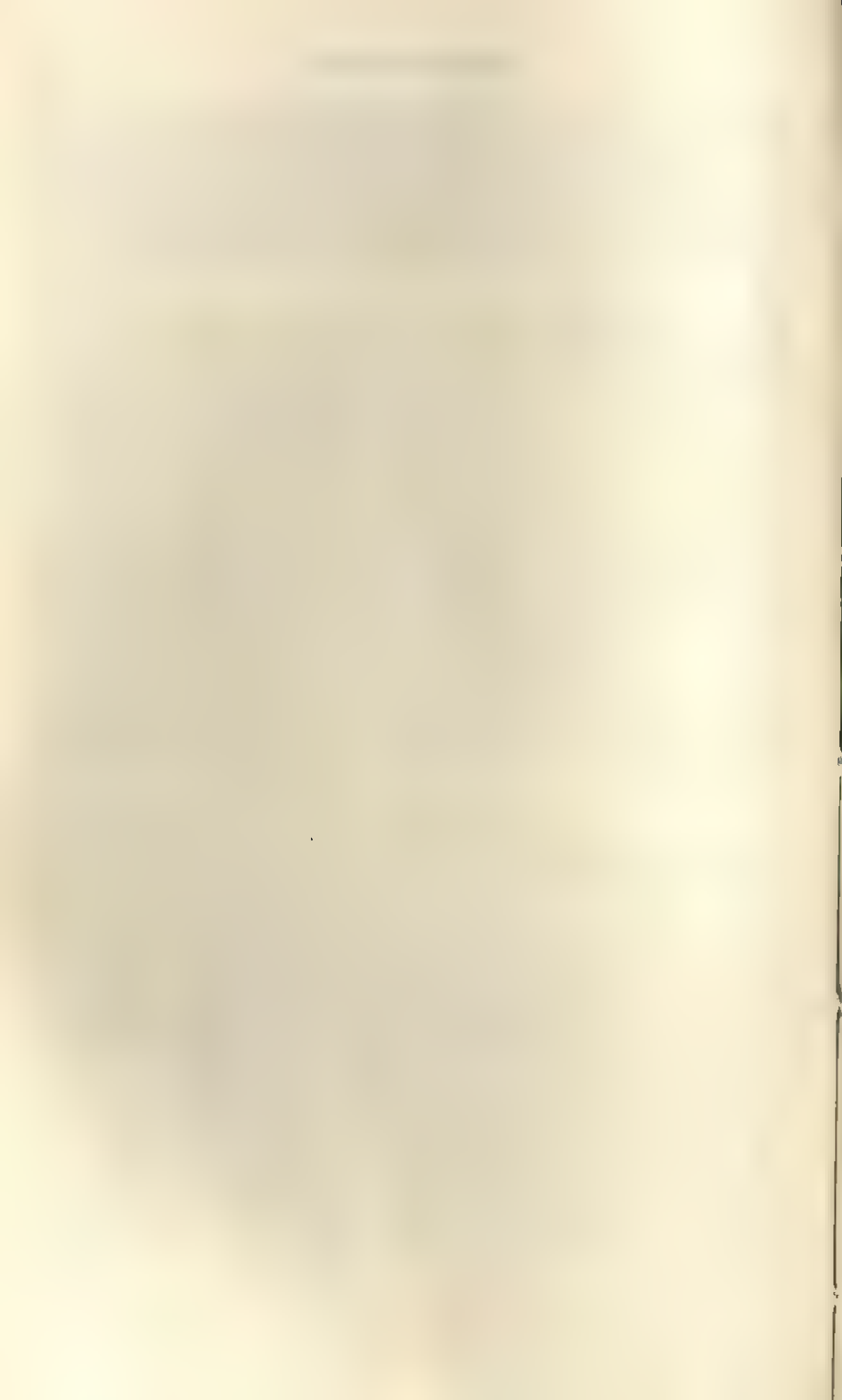
#### REFERENCES

1. BOULDING, K. *The Organizational Revolution*. New York: Harper & Row, 1953.
2. BRUNER, J. S., GOODNOW, J. J., & AUSTIN, G. A. *A Study of Thinking*. New York: Wiley, 1956.
3. CLARK, R. D., & WILLEMS, E. P. Risk preferences as related to judged consequences of failure. *Psychol. Rep.*, 1969, 25, 827-830.
4. FLANDERS, J. P. Generality of the risky shift phenomenon to a task with demonstrably nontrivial decision consequences. Unpublished manuscript, Walter Reed Army Institute of Research, Washington D. C., 1970.
5. JOHNSON, D. L., & ANDREWS, R. L. Risky-shift phenomenon tested with consumer products as stimuli. *J. Personal. & Soc. Psychol.*, 1971, 20, 382-385.
6. KOGAN, N., & WALLACH, M. A. *Risk Taking: A Study in Cognition and Personality*. New York: Holt, Rinehart & Winston, 1964.
7. ———. Risk taking as a function of the situation, the person, and the group. G. Mandler, P. Mussen, N. Kogan, & M. Wallach (Eds.), *New Directions in Psychology III*. New York: Holt, Rinehart & Winston, 1967.
8. MACKENZIE, K. D. An analysis of risky shift experiments. *Organiz. Behav. & Hum. Perform.*, 1971, 6, 283-303.
9. RICHMOND, S. B. *Operations Research for Management Decisions*. New York: Ronald Press, 1968.

10. SIMON, H. A. *Models of Man*. New York: Wiley, 1957.
11. STONER, J. A. F. A comparison of individual and group decisions involving risk. Unpublished Master's thesis, Massachusetts Institute of Technology, Sloan School of Management, Cambridge, Massachusetts, 1961.
12. ———. Risky and cautious shifts in group decisions: The influence of widely held values. *J. Exper. Soc. Psychol.*, 1968, 4, 442-459.
13. WHYTE, W. H., JR. *The Organization Man*. New York: Simon & Schuster, 1956.

*Department of Marketing*  
*College of Business Administration*  
*University of Cincinnati*  
*Cincinnati, Ohio 45221*





# The Journal of PSYCHOLOGY

## The General Field of Psychology

POWELL MURCHISON  
*Managing Editor*

HELEN M. DAVIS  
*Associate Editor*

### *Editorial Board*

JOHN A. ANTOINETTI  
Life Insurance Agency  
Management Association

S. HOWARD BARTLEY  
Michigan State University

ARTHUR G. BILLS  
University of Cincinnati

E. A. BOTT  
University of Toronto

LEONARD CARMICHAEL  
National Geographic Society

IRVIN L. CHILD  
Yale University

HERBERT S. CONRAD  
U. S. Office of Education

JOHN F. DASHIELL  
Florida State University

LEONARD W. DOOB  
Yale University

If this space should be unstamped, this is the regular library edition.  
But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not  
be resold for a period of five years from date of publication.

FRANK A. GELDARD  
Princeton University

ERNEST R. HILGARD  
Stanford University

JOHN E. HORROCKS  
Ohio State University

DAVID HOTHERSALL  
Ohio State University

HEINRICH KLÜVER  
University of Chicago

NORMAN L. MUNN  
University of Adelaide,  
Australia

WARREN H. TEICHNER  
New Mexico State  
University

LEE EDWARD TRAVIS  
Fuller Theological  
Seminary

JOSEPH ZUBIN  
Psychiatric Institute,  
Columbia University

## VOLUME 85

1 9 7 3

Copyright © 1973 by the Journal Press  
Provincetown, Massachusetts, U. S. A.

All rights reserved

Published at 2 Commercial Street, Provincetown, Mass.

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

# AUTHOR INDEX

Bailey, Kent G. ....	143	Lerner, Richard M. ....	119
Beck, Aaron T. ....	257	Lester, David ....	257
Bronzajt, Arline L. ....	17	Mahoney, John ....	37
Brown, Booker ....	43	Mangelsdorff, A. David ....	249
Burdsal, Charles ....	45	McCall, James ....	87
Cull, John G. ....	225	McCarthy, Kevin ....	53
Dean, Michael L. ....	245	McKinney, John Paul ....	235
Dunlap, Sally M. ....	249	Menges, Robert J. ....	69
Dwyer, Francis M. ....	297	Mills, John A. ....	313
Edgington, Eugene S. ....	89	Minogue, Ellen M. ....	301
Epstein, Gilda F. ....	17	Monacelli, Vincent ....	75
Forehand, Rex ....	323	Nedd, Albert N. B. ....	305
Gaertner, Samuel L. ....	249	O'Dell, Jerry W. ....	165
Gardiner, Harry W. ....	97	Pedersen, Darhl M. ....	101
Gardner, Harold L. ....	323	Pellegrini, Robert J. ....	21
Gerbasi, Kathleen Carrese ....	109	Poe, Charles A. ....	81
Gilpin, Andrew R. ....	277	Rae, Gordon ....	87
Glover, Hilda ....	143	Reingen, Peter H. ....	339
Goldblatt, Robert B. ....	75	Rhoads, Caryln S. ....	329
Goldman, Morton ....	329	Riess, Bernard F. ....	61
Greenberg, Gary ....	45	Rutschmann, Jacques ....	187
Haddad, Jay ....	53	Sack, Stanley ....	3
Hardy, Robert C. ....	3	Safer, Jeanne ....	61
Harpine, Frances ....	3	Schaefer, Charles ....	179
Hartnett, Jack ....	37	Sears, Robert R. ....	267
Hartnett, John J. ....	143	Sewell, Alan F. ....	151, 157
Heacock, Del ....	43	Smith, Otho H. ....	225
Heisler, James T. ....	151	Steininger, Marion ....	131
Johnson, Belinda ....	69	Strahan, Robert ....	109
Karabenick, Stuart A. ....	119	Strain, Allan R. ....	89
Kelleher, John H. ....	261	Strumpfer, D. J. W. ....	29
Kling, John O. ....	301	Stuart, Joyce L. ....	119
Koslowsky, Meni ....	287	Thurber, Steven ....	43
Lackner, James R. ....	137	Timpe, Randie ....	45
La Fave, Lawrence ....	53	Very, Philip S. ....	75
Laird, J. T. ....	171	Way, John G. ....	229
Lapidus, Deborah ....	267	Williams, John D. ....	261

# TABLE OF CONTENTS

An experimental test of the contingency model on small classroom groups . . . . .	3
By ROBERT C. HARDY, STANLEY SACK, AND FRANCES HARPINE	
Is test anxiety rising? . . . . .	17
By ARLINE L. BRONZAIT AND GILDA F. EPSTEIN	
The astrological "theory" of personality: An unbiased test by a biased observer . . . . .	21
By ROBERT J. PELLEGRINI	
Failure to find relationships between family constellation and achievement motivation . . . . .	29
By D. J. W. STRÜMPFER	
Self-actualization and self-ideal discrepancy . . . . .	37
By JOHN MAHONEY AND JACK HARTNETT	
The control orientation of black athletes in relation to unobtrusive tasks of skill . . . . .	43
By STEVEN THURBER, DEL HEACOCK, AND BOOKER BROWN	
The relationship of marihuana usage to personality and motivational factors . . . . .	45
By CHARLES BURDSAL, GARY GREENBERG, AND RANDIE TIMPE	
Humor judgments as a function of identification classes: Canadian <i>vs.</i> American . . . . .	53
By LAWRENCE LA FAVE, KEVIN MCCARTHY, AND JAY HADDAD	
Birth order and related variables in a large outpatient population . . . . .	61
By BERNARD F. RIESS AND JEANNE SAFER	
Implicit time constraints in the measurement of productive thinking . . . . .	69
By ROBERT J. MENGES AND BELINDA JOHNSON	
Birth order, personality development, and vocational choice of becoming a Carmelite nun . . . . .	75
By PHILIP S. VERY, ROBERT B. GOLDBLATT, AND VINCENT MONACELLI	
Development of a psychological effectiveness scale . . . . .	81
By CHARLES A. POE	
Some international comparisons of cancer mortality rates and personality: A brief note . . . . .	87
By GORDON RAE AND JAMES MCCALL	
Randomization tests: Computer time requirements . . . . .	89
By EUGENE S. EDGINGTON AND ALLAN R. STRAIN	
Catholic sisters and the Edwards Personal Preference Schedule . . . . .	97
By HARRY W. GARDINER	
Personality and demographic correlates of simulated personal space . . . . .	101
By DARHL M. PEDERSEN	
Semantic style variance in personality questionnaires . . . . .	109
By ROBERT STRAHAN AND KATHLEEN CARRESE GERBASI	

Relations among physical attractiveness, body attitudes, and self-concept in male and female college students . . . . .	119
BY RICHARD M. LERNER, STUART A. KARABENICK, AND JOYCE L. STUART	
In defense of measuring attitudes . . . . .	131
BY MARION STEININGER	
A device for investigating adaptation to sensory rearrangement . . . . .	137
BY JAMES R. LACKNER	
Modeling and personal space behavior in children . . . . .	143
BY KENT G. BAILEY, JOHN J. HARTNETT, AND HILDA W. GLOVER	
Personality correlates of proximity preferences . . . . .	151
BY ALAN F. SEWELL AND JAMES T. HEISLER	
Person perception as a function of the personal consequences and immediacy of a decision . . . . .	157
BY ALAN F. SEWELL	
The EPSAT as a predictor . . . . .	165
BY JERRY W. O'DELL	
Mental health and population density . . . . .	171
BY J. T. LAIRD	
Measuring social relationships in emotionally disturbed boys . . . . .	179
BY CHARLES SCHAEFER	
Time judgments by magnitude estimation and magnitude production and anxiety: A problem of comparison between normals and certain schizophrenic patients . . . . .	187
BY JACQUES RUTSCHMANN	
A preliminary note on demographic and personality correlates of decubitus ulcer incidence . . . . .	225
BY JOHN G. CULL AND OTHO H. SMITH	
The relation of student and teacher traits of authoritarianism to student achievement in English . . . . .	229
BY JOHN G. WAY	
The structure of behavioral values of college students . . . . .	235
BY JOHN PAUL MCKINNEY	
The impact of exam question order effects on student evaluations . . . . .	245
BY MICHAEL L. DEAN	
Personality and attitudes: A re-emphasis upon the cognitive component . . . . .	249
BY SALLY M. DUNLAP, SAMUEL L. GAERTNER, AND A. DAVID MANGELSDORFF	
Components of depression in attempted suicides . . . . .	257
BY AARON T. BECK AND DAVID LESTER	
Conceptual systems and philosophical orientation . . . . .	261
BY JOHN D. WILLIAMS AND JOHN H. KELLEHER	
Episodic analysis of novels . . . . .	267
BY ROBERT R. SEARS AND DEBORAH LAPIDUS	



Lexical marking effects in the semantic differential . . . . .	277
BY ANDREW R. GILPIN	
The stereotypes of four ethnic groups . . . . .	287
BY MENI KOSLOWSKY	
The relative effectiveness of two methods of presenting visualized instruction . . . . .	297
BY FRANCIS M. DWYER	
Massed <i>versus</i> spaced sessions in systematic desensitization . . . . .	301
BY JOHN O. KLING AND ELLEN M. MINOGUE	
Managerial attitudes toward hiring ex-convicts . . . . .	305
BY ALBERT N. B. NEDD	
Effects of interitem associative strength, rehearsal, and proactive inhibition on the retention of paired-associate lists . . . . .	313
BY JOHN A. MILLS	
An examination of verbal imitative performance in young children . . . . .	323
BY REX FOREHAND AND HAROLD L. GARDNER	
Frustration response categories and level of hostile expression . . . . .	329
BY MORTON GOLDMAN AND CARYLN S. RHODES	
Risk taking by individuals and informal groups with the use of industrial product purchasing situations as stimuli . . . . .	339
BY PETER H. REINGEN	

U. S. POSTAL SERVICE <b>STATEMENT OF OWNERSHIP, MANAGEMENT AND CIRCULATION</b> <small>(Act of August 12, 1970. Section 3685, Title 39, United States Code)</small>		SEE INSTRUCTIONS ON PAGE 2 (REVERSE)
1. TITLE OF PUBLICATION <b>Journal of Psychology</b>		2. DATE OF FILING <b>Sept. 30, 1973</b>
3. FREQUENCY OF ISSUE <b>6 issues; published Jan., March, May, July, Sept., Nov.</b>		
4. LOCATION OF KNOWN OFFICE OF PUBLICATION (Street, city, county, state, ZIP code) (Not printers) <b>The Journal Press, 2 Commercial St., Provincetown, Barnstable, Ma. 02657</b>		
5. LOCATION OF THE HEADQUARTERS OR GENERAL BUSINESS OFFICES OF THE PUBLISHERS (Not printers) <b>The Journal Press, 2 Commercial Street, Provincetown, Mass. 02657</b>		
6. NAMES AND ADDRESSES OF PUBLISHER, EDITOR, AND MANAGING EDITOR		
PUBLISHER (Name and address) <b>The Journal Press, 2 Commercial Street, Provincetown, Mass. 02657</b>		
EDITOR (Name and address) <b>Editorial Board, 2 Commercial Street, Provincetown, Mass. 02657</b>		
MANAGING EDITOR (Name and address) <b>Powell Murchison, 2 Commercial Street, Provincetown, Mass. 02657</b>		
7. OWNER (If owned by a corporation, its name and address must be stated and also immediately thereunder the names and addresses of stockholders owning or holding 1 percent or more of total amount of stock. If not owned by a corporation, the names and addresses of the individual owners must be given. If owned by a partnership or other unincorporated firm, its name and address, as well as that of each individual must be given.)		
NAME	ADDRESS	
<b>The Journal Press</b>	<b>2 Commercial, Provincetown, Ma. 02657</b>	
<b>Dorotea P. Murchison</b>	<b>2 Commercial, Provincetown, Ma. 02657</b>	
<b>Powell Murchison</b>	<b>2 Commercial, Provincetown, Ma. 02657</b>	
<b>Marjorie E. Monteth</b>	<b>Mayflower Hts., Provincetown, Ma. 02657</b>	
8. KNOWN BONDHOLDERS, MORTGAGEES, AND OTHER SECURITY HOLDERS OWNING OR HOLDING 1 PERCENT OR MORE OF TOTAL AMOUNT OF BONDS, MORTGAGES OR OTHER SECURITIES (If there are none, so state)		
NAME	ADDRESS	
<b>Same as above</b>		
9. FOR OPTIONAL COMPLETION BY PUBLISHERS MAILING AT THE REGULAR RATES (Section 132.121, Postal Service Manual)		
<p>39 U. S. C. 3626 provides in pertinent part: "No person who would have been entitled to mail matter under former section 4359 of this title shall mail such matter at the rates provided under this subsection unless he files annually with the Postal Service a written request for permission to mail matter at such rates."</p> <p>In accordance with the provisions of this statute, I hereby request permission to mail the publication named in item 1 at the reduced postage rates presently authorized by 39 U. S. C. 3626.</p> <p>(Signature and title of editor, publisher, business manager, or owner) <b>Powell Murchison, Business Manager</b></p>		
10. FOR COMPLETION BY NONPROFIT ORGANIZATIONS AUTHORIZED TO MAIL AT SPECIAL RATES (Section 132.122, Postal Manual)		
(Check one)		
The purpose, function, and nonprofit status of this organization and the exempt status for Federal income tax purposes	<input type="checkbox"/> Have not changed during preceding 12 months	<input type="checkbox"/> Have changed during preceding 12 months (If changed, publisher must submit explanation of change with this statement.)
11. EXTENT AND NATURE OF CIRCULATION		
A. TOTAL NO. COPIES PRINTED (Net Press Run)	2225	2225
B. PAID CIRCULATION		
1. SALES THROUGH DEALERS AND CARRIERS, STREET VENDORS AND COUNTER SALES	none	none
2. MAIL SUBSCRIPTIONS	1690	1717
C. TOTAL PAID CIRCULATION	1690	1717
D. FREE DISTRIBUTION BY MAIL, CARRIER OR OTHER MEANS		
1. SAMPLES, COMPLIMENTARY, AND OTHER FREE COPIES	26	26
2. COPIES DISTRIBUTED TO NEWS AGENTS, BUT NOT SOLD	none	none
E. TOTAL DISTRIBUTION (Sum of C and D)	1716	1743
F. OFFICE USE, LEFT-OVER, UNACCOUNTED, SPOILED AFTER PRINTING	509	482
G. TOTAL (Sum of E & F—should equal net press run shown in A)	2225	2225
(Signature of editor, publisher, business manager, or owner) <b>Powell Murchison, Business Manager</b>		
I certify that the statements made by me above are correct and complete.		

# ABBREVIATIONS OF WORDS USED IN THE TITLES OF JOURNALS

(One-word titles are never abbreviated. The word "the" is not used, nor its equivalent in any other language. The word "of" or its equivalent in other languages is used only to discriminate what would otherwise be identical titles in different languages. The word "and" is always used, but indicated by "&" in the Roman alphabet. Only English words are indicated here, but the corresponding words in other languages should receive a corresponding abbreviation. All abbreviations and all one-word titles should be in italics.)

Abnormal .....	<i>Abn.</i>	Japanese .....	<i>Jap.</i>
Abstracts .....	<i>Abst.</i>	Journal .....	<i>J.</i>
American .....	<i>Amer.</i>	Mathematical .....	<i>Math.</i>
Anatomy .....	<i>Anat.</i>	Measurement .....	<i>Meas.</i>
Animal .....	<i>Anim.</i>	Medical .....	<i>Med.</i>
Applied .....	<i>Appl.</i>	Mental .....	<i>Ment.</i>
Archives .....	<i>Arch.</i>	Monographs .....	<i>Monog.</i>
Association .....	<i>Assoc.</i>	Neurology .....	<i>Neurol.</i>
Attitude .....	<i>Attit.</i>	Opinion .....	<i>Opin.</i>
Australian .....	<i>Aust.</i>	Orthopsychiatry .....	<i>Orthopsychiat.</i>
Behavior .....	<i>Behav.</i>	Personality .....	<i>Personal.</i>
British .....	<i>Brit.</i>	Personnel .....	<i>Person.</i>
Bulletin .....	<i>Bull.</i>	Philosophy .....	<i>Philos.</i>
Bureau .....	<i>Bur.</i>	Physics .....	<i>Phys.</i>
Canadian .....	<i>Can.</i>	Physiology .....	<i>Physiol.</i>
Character .....	<i>Charac.</i>	Proceedings .....	<i>Proc.</i>
Children .....	<i>Child.</i>	Psychiatry .....	<i>Psychiat.</i>
Chinese .....	<i>Chin.</i>	Psychoanalysis .....	<i>Psychoanal.</i>
Clinical .....	<i>Clin.</i>	Psychology .....	<i>Psychol.</i>
College .....	<i>Coll.</i>	Psychosomatic .....	<i>Psychosomat.</i>
Comparative .....	<i>Comp.</i>	Quarterly .....	<i>Quart.</i>
Consulting .....	<i>Consult.</i>	Religious .....	<i>Relig.</i>
Contributions .....	<i>Contrib.</i>	Research .....	<i>Res.</i>
Development .....	<i>Devel.</i>	Review .....	<i>Rev.</i>
Educational .....	<i>Educ.</i>	School .....	<i>Sch.</i>
Experimental .....	<i>Exper.</i>	Science .....	<i>Sci.</i>
General .....	<i>Gen.</i>	Social .....	<i>Soc.</i>
Genetic .....	<i>Genet.</i>	Statistics .....	<i>Stat.</i>
Indian .....	<i>Ind.</i>	Studies .....	<i>Stud.</i>
Industrial .....	<i>Indus.</i>	Teacher .....	<i>Teach.</i>
International .....	<i>Internat.</i>	University .....	<i>Univ.</i>
Italian .....	<i>Ital.</i>		

9.7.74

# Preparation of Manuscripts for The Journal Press

## GENERAL INSTRUCTIONS

1. The proper sequence for the parts of your submitted manuscript is as follows: (a) text, (b) references, (c) footnotes, (d) tables, (e) figures, and (f) figure legends. However, monographs start with a table of contents and may have an acknowledgment page before the text and an appendix immediately after the text.
2. Use heavy typewriter paper,  $8\frac{1}{2} \times 11$  inches, double-space *all* lines, and leave margins for editorial work. Do not use onionskin, odd sizes, and abrasive or wax finishes.
3. Submit original typewritten version and one copy. Retain second copy for proofing.
4. Retype any page on which written corrections have been made.
5. Do not begin a sentence with a numeral.
6. A summary at the beginning of the text is required for articles over 500 words.
7. Each quotation should indicate the page number of the original source. The original publisher must give permission for lengthy quotations and use of tables or figures.
8. Do not fold your manuscript.
9. Enclose a submission letter, with a statement that the manuscript is not under consideration elsewhere. If you are unknown to the Editors, kindly give your credentials.

## FORMAT AND SPECIFIC INSTRUCTIONS

### A. TEXT DIVISIONS

#### I. THE TITLES OF JOURNAL ARTICLES AND THE MAJOR SUBDIVISIONS OF MONOGRAPHS ARE PRINTED IN TEN-POINT CAPS CENTERED ON THE PAGE

##### A. THE NEXT SUBDIVISION TITLE IS PRINTED IN CAPS AND SMALL CAPS CENTERED ON THE PAGE

1. Then *Italics, with Principal Words, Upper and Lower Case, Centered on the Page*
- a. Then *italics, upper and lower case, 1-em run-in side head.*

(1). Then *italics, upper and lower case, 2-em run-in side head.*

(a). Then *italics, upper and lower case, 3-em run-in side head.*

[Further subdivision should be merged into the text without marginal indentation, and should be numbered with small letters.]

### B. REFERENCES

References should be arranged in alphabetical order by author, numbered and referred to in the text by number (2). Double-space!

The proper form of a book reference is as follows:

2. DOB, J. The Preparation of Manuscripts. New York: Holt, 1963. Pp. 400-418.

The proper form of a journal reference is as follows:

2. DOB, J. The preparation of manuscripts. *J. Gen. Psychol.*, 1963, 68, 450-462.

If in the text it is desirable to refer to a page, thus (2, p. 45).

### C. FOOTNOTES

Use as few as possible and number consecutively in the text thus.<sup>1</sup>

<sup>1</sup> Footnote (on the separate footnote page). Double-space!

### D. TABLES

Each table should be typed on a separate sheet and should be Arabic numbered (Table 2). Each column requires a heading. Vertical lines should be avoided. Mention tables consecutively in the text and indicate approximate insertion points.

### E. FIGURES

Figures should be submitted as glossy prints of the approximate size for final reproduction (ordinarily  $4\frac{1}{2}$  inches in width). Lettering and lines should be sharp and clean. Figures should be Arabic numbered (Figure 2) in the text, and have consecutive mention and approximate insertion points. Each figure requires a legend, but all legends should be submitted double-spaced on a separate Figure Legends page.

\$39.00 per annum  
\$18.00 per volume  
Single Numbers \$9.00

BIMONTHLY  
Three volumes per year  
Immediate publication

September, 1973  
Volume 85, First Half

Founded in 1935 by Carl Murchison

# The Journal of PSYCHOLOGY

The General Field of Psychology

POWELL MURCHISON  
*Managing Editor*

HELEN M. DAVIS  
*Associate Editor*

## Editorial Board

JOHN A. ANTOINETTI  
Life Insurance Agency  
Management Association

S. HOWARD BARTLEY  
Michigan State University

ARTHUR G. BILLS  
University of Cincinnati

E. A. BOTT  
University of Toronto

LEONARD CARMICHAEL  
National Geographic Society

IRVIN L. CHILD  
Yale University

HERBERT S. CONRAD  
U. S. Office of Education

JOHN F. DASHIELL  
Florida State University

LEONARD W. DOOB  
Yale University

If this space should be unstamped, this is the regular library edition.  
But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not  
be resold for a period of five years from date of publication.

FRANK A. GELDARD  
Princeton University

ERNEST R. HILGARD  
Stanford University

JOHN E. HORROCKS  
Ohio State University

DAVID HOTHERSALL  
Ohio State University

HEINRICH KLÜVER  
University of Chicago

NORMAN L. MUNN  
University of Adelaide,  
Australia

WARREN H. TRICHTER  
New Mexico State  
University

LEE EDWARD TRAVIS  
Fuller Theological  
Seminary

JOSEPH ZUBIN  
Psychiatric Institute,  
Columbia University

Copyright © 1973 by The Journal Press  
Provincetown, Massachusetts, U. S. A.

All rights reserved

Second-class postage paid at Provincetown, Mass.  
and additional mailing offices

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657



# Microfiche Now Available

## Back Issues

While continuing to maintain a complete inventory of paper-bound back issues for all five journals—some reprinted and some original editions—The Journal Press is now offering microfiche (negative silverfiche) versions for all but the last two calendar years: i.e., Volume 1 through 1970 volumes are available immediately; 1971 volumes will first be available in January 1974; 1972 volumes, in January 1975, etc. The minimum order is for one volume. The following rates are in effect: Paper-bound—single numbers, \$9.00; microfiche, not available. Paper-bound—two-part volumes, \$18.00; microfiche, \$6.00. Paper-bound—three-part volumes, \$27.00; microfiche, \$9.00. Paper-bound—four-part and over-four-part volumes, \$36.00; microfiche, \$12.00. The only exceptions (special issues) are starred below.\*

VOLUMES	DATES	PRICE PER VOLUME		PRICE THROUGH 1970	
		PAPER	MICRO-FICHE	PAPER	MICRO-FICHE
<i>Journal of Genetic Psychology</i> (Pedagogical Seminary)					
1-117	1891-1970			\$2754.00	\$895.00
1-4	1891-April 1897	\$27.00	\$ 9.00		
5-37	July 1897-1930	36.00	12.00		
38*	1930	18.00	6.00		
39	1931	36.00	12.00		
40 and later	1932-	18.00	6.00		
<i>Genetic Psychology Monographs</i>					
1-82	1926-1970			1872.00	620.00
1-22	1926-1940	36.00	12.00		
23 and later	1941-	18.00	6.00		
<i>Journal of General Psychology</i>					
1-83	1928-1970			1566.00	520.00
1-3	1928-1930	36.00	12.00		
4*	1930	18.00	6.00		
5	1931	36.00	12.00		
6 and later	1932-	18.00	6.00		
<i>Journal of Social Psychology</i>					
1-82	1930-1970			1656.00	550.00
1-10	1930-1939	36.00	12.00		
11 and later	1940-	18.00	6.00		
<i>Journal of Psychology</i>					
1-76	1935/36-1970			1368.00	450.00
1 and later	1935/36-	18.00	6.00		

A mailing charge of \$0.30 per volume is made for foreign delivery of paper editions. Domestic paper mailings carry no postage charge. Prepaid foreign and domestic microfiche orders are also postage free; other orders will be billed for costs. Remittances should be by check or money order payable in U.S. funds on a United States bank, with no bank commission deducted. Advance payment is required from all individuals and from foreign purchasers. Microfilm versions and special bindings are not available. Microfiche orders only may be sent directly to The Journal Press or to Johnson Associates Inc., 175 Fifth Avenue, New York, New York 10010.

## THE JOURNAL PRESS

2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.

*Notice: All copyrighted material reprinted or otherwise reproduced from Journal Press publications should be given an appropriate acknowledgment. No permission is necessary for brief quotations and originally worded abstracts (less than 300-word total or 10 percent of length, whichever is least, for any one article). Written permission is required for reprinting or copying tables, figures, summaries, entire articles, or substantial portions thereof; such permission normally carries a charge of \$5 per page and is contingent upon agreement and possible stipulation of the original author or authors. Estate permission can be required. Authors requesting use of their own material are not charged. Fees may be waived for limited-copy instructional usage and for government periodicals. Special rates can be applied when justified.*

# The Journal of Psychology

Founded in 1935 by Carl Murchison

The Journal of Psychology was founded at the request of many psychologists throughout the world for the purpose of establishing a medium in which publication shall be immediate and in which other publication experiments may be carried out.

(1). The subscription price is \$39.00 per annum, \$18.00 per volume, \$9.00 per issue (three volumes per year).

(2). Manuscripts may be sent to any member of the Editorial Board or may be sent directly to The Journal Press. The lag in publication should not be more than a few weeks.

(3). The author will receive 200 reprints of his paper, and subscribers will receive the accumulated publications bound into half-volume units. The second half of each volume will contain the title page and table of contents for that volume. Each half-volume unit will be mailed to the subscribers at its stated date.

(4). Each author will be charged the manufacturing costs of any zincs, halftone cuts, tables, or equations. Excessive use of special fonts (references, footnotes, quotations, appendices) will also be charged.

(5). The publication charges are as follows:

1 to 4 pages	\$60.00
5 to 8 "	120.00
9 to 12 "	180.00
13 to 16 "	240.00
17 to 20 "	300.00
21 to 24 "	360.00
25 to 28 "	420.00
29 to 32 "	480.00

These charges are subject to adjustment as justified.

(6). The accumulated units will be mailed to subscribers bimonthly: January, March, May, July, September, and November. The units may vary in size to accommodate publication needs.

All back issues are available. See announcement on inside front cover.

THE JOURNAL PRESS  
2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.

Diary No. 696  
Date 12-12-7  
46

# Journal of Genetic Psychology

Founded by G. Stanley Hall in 1891; reorganized by Carl Murchison in 1925

Quarterly, beginning with March. Devoted to developmental and clinical psychology, with briefly reported Replications and Refinements and occasional book reviews. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$26.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

---

# Journal of General Psychology

Founded in 1927 by E. B. Titchener and Carl Murchison

Quarterly, beginning with January. Devoted to experimental, physiological, and comparative psychology, with briefly reported Replications, Refinements, and Comments. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$26.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

---

# Journal of Social Psychology

Founded in 1929 by John Dewey and Carl Murchison

Bimonthly, beginning with February. Devoted to studies of persons in group settings and of culture and personality; special attention to cross-cultural articles and notes, to field research, and to briefly reported Replications and Refinements. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Three volumes annually. Per annum, \$39.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

---

*Each author will receive 100 unbound reprints gratis. Additional reprints may be purchased at a nominal cost when the author returns page-proof.*

THE JOURNAL PRESS

2 Commercial Street

Provincetown, Massachusetts 02657

U. S. A.

393

• 51 • 51

051

# Club Rate for Five International Psychological Journals

	PER ANNUM
THE JOURNAL OF GENETIC PSYCHOLOGY . . . . . (Quarterly, two volumes per year)	\$26.00
GENETIC PSYCHOLOGY MONOGRAPHS . . . . . (Quarterly, two volumes per year)	\$26.00
THE JOURNAL OF GENERAL PSYCHOLOGY . . . . . (Quarterly, two volumes per year)	\$26.00
THE JOURNAL OF SOCIAL PSYCHOLOGY . . . . . (Bimonthly, three volumes per year)	\$39.00
THE JOURNAL OF PSYCHOLOGY . . . . . (Bimonthly, three volumes per year)	\$39.00
TOTAL	\$156.00

TEACHERS AND STUDENT'S EDITION, *Restricted Sale* \$50.00

The *Teachers and Students Edition* is open to any individual psychologist, teacher, student, or physician but requires an agreement that the edition will not be offered for sale till after five years from publication and an agreement that the edition will not be substituted for an institutional subscription. Write for our Teachers and Students contract card.

## THE JOURNAL PRESS

2 Commercial Street  
Provincetown, Massachusetts 02657  
U. S. A.

## Genetic Psychology Monographs

Founded in 1925 by Carl Murchison

Quarterly, beginning with February. Devoted to developmental and clinical psychology. Each number one or more complete researches. Manuscripts may be sent to any member of the Editorial Board or directly to The Journal Press. Subscriptions should be sent directly to The Journal Press, and business communications addressed to the Managing Editor. Two volumes annually. Per annum, \$26.00; per volume, \$18.00; single numbers, \$9.00. All back issues available. See announcement on inside front cover.

**NOTICE.** Effective immediately we offer prior publication for certain selected monographs. Charges will be based on the announced rates for *The Journal of Psychology*. Regular monographs will carry no page charge for the first 50 pages, but for all pages above 50 there will be a charge of \$10.00 per page. Manufacturing costs of any zincs, halftone cuts, tables, or equations will be charged to authors, as will excessive use of special fonts. Two hundred reprints of each monograph will be supplied free of charge.

The listing of monographs for 1926 through 1963 will be printed in the first half of each volume of all five journals. Monographs published after 1963 will be listed in the second half of each volume.

# Genetic Psychology Monographs (continued)

## VOLUME 1—1926

1. Performance tests for children of pre-school age—R. SPITSMAN
2. An experimental study of the eidetic type—H. KLÜVER
- 3 & 4. A study of natio-racial mental differences—N. D. M. HIRSCH
5. A psychological study of juvenile delinquency by group methods—J. W. BRIDGES AND K. M. B. BRIDGES
6. The influence of puberty praecox upon mental growth—A. GSEL

## VOLUME 2—1927

- 1 & 2. The mind of a gorilla—R. M. YERKES
3. The role of eye-muscles and mouth-muscles in the expression of the emotions—K. DUNLAP
4. Family similarities in mental-test abilities—R. A. WILLOUGHBY
5. Coordination in the locomotion of infants—L. H. BRANSIDE
6. The mind of a gorilla: Part II. Mental development—R. M. YERKES

## VOLUME 3—January-June, 1928

1. An experimental study of the olfactory sensitivity of the white rat—J. R. LIGGETT
2. A photographic study of eye movements in reading formulae—M. A. TINKER
3. An experimental study of the East Kentucky mountaineers—N. D. M. HIRSCH
4. Responses of foetal guinea pigs prematurely delivered—G. T. AVERY
5. Objective differentiation between three groups in education (teachers, research workers, and administrators)—M. B. JENSEN
6. The effect of segregation on the sex behavior of the white rat as measured by the abstraction method—M. JENSEN

## VOLUME 4—July-December, 1928

1. Observation and training of fundamental habits in young children—E. A. BOTT, W. F. BLATZ, N. CHANT, AND H. BOTT
- 2 & 3. Determination of a content of the course in literature of a suitable difficulty for junior and senior high school students—M. C. BURCH
- 1 & 5. Methods for diagnosis and treatment of cases of reading disability—M. MOSKOW
6. The relative effectiveness of lecture and individual reading as methods of college teaching—E. B. CHURCH

## VOLUME 5—January-June, 1929

1. The age factor in animal learning: I. Rats in the problem box and the maze—C. P. STONE
2. The effect of delayed incentive on the hunger drive in the white rat—E. L. HAMILTON
3. Which hand is the eye of the blind?—J. M. SMITH
4. The effect of attitude on free word association-time—A. G. ECKHAUL
5. The localization of tactual space: A study of average and constant errors under different types of localization—L. E. COLE
6. The effects of gonadectomy, vasotomy, and injections of placental and orchic extracts on the sex behavior of the white rat—H. W. NISSEN

## VOLUME 6—July-December, 1929

1. Learning and growth in identical infant twins: An experimental study by the method of co-twin control—A. GSEL AND H. THOMPSON
2. The age factor in animal learning: II. Rats on a multiple light discrimination box and a difficult maze—C. P. STONE
3. The acquisition and interference of motor habits in young children—E. MCGINNIS
4. A vocational and socio-educational survey of graduates and non-graduates of small high schools of New England—A. D. MITCHELL
- 5 & 6. A study of the smiling and laughing of infants in the first year of life—R. W. WASHBURN

## VOLUME 7—January-June, 1930

1. Tensions and emotional factors in reaction—E. DUFFY
2. Teacher influence on class achievement: A study of the relationship of estimated teaching ability to pupil achievement in reading and arithmetic—H. R. TAYLOR
- 3 & 4. A study of the effect of inverted retinal stimulation upon spatially coordinated behavior—P. H. EMMETT
5. A study of the mental development of children with lesion in the central nervous system—F. E. LOEB
6. An experimental study upon three hundred school children over a six-year period—N. D. M. HIRSCH

## VOLUME 8—July-December, 1930

1. The amount and nature of activities of newborn infants under constant external stimulating conditions during the first ten days of life—O. C. LAWIN
2. Race and social differences in performance tests—S. D. PORTEUS, *et al.*
3. Language and growth: The relative efficacy of early and deferred vocabulary training, studied by the method of co-twin control—L. C. STRAYER
4. Eye-movements and optic nystagmus in early infancy—J. M. MCGINNIS
- 5 & 6. Reactions of kindergarten, first-, and second-grade children to constructive play materials—L. FARWELL

## VOLUME 9—January-June, 1931

- 1 & 2. The status of the first-born with special reference to intelligence—H. H. HEBB
- 3 & 4. An experimental study of bright, average, and dull children at the four-year mental level—H. P. DAVIDSON
5. An historical, critical, and experimental study of the Seashore-Kwalwasser test battery—P. R. FARNWORTH
6. A comparison of difficulty and improvement in the learning of bright and dull children in reproducing a descriptive selection—F. T. WILSON

## VOLUME 10—July-December, 1931

1. A comparative study of a group of southern white and negro infants—M. B. MCGRAW
- 2 & 3. An experimental study of prehension in infants by means of systematic cinema records—H. M. HALVENSEN
4. The limits of learning ability in kittens—A. M. SHUEY
- 5 & 6. The effect of habit interference upon performance in maze learning—O. W. ALM

## VOLUME 11—January-June, 1932

1. General factors in transfer of training in the white rat—T. A. JACKSON
2. The effect of color on visual apprehension and perception—M. A. TINKER
3. The reliability and validity of maze experiments with white rats—R. LEEPER
4. A critical study of two lists of best books for children—F. K. SHUTTLEWORTH
- 5 & 6. Measuring human energy cost in industry: A general guide to the literature—R. M. PAGE

## VOLUME 12—July-December, 1932

1. Family resemblances in verbal and numerical abilities—H. D. CARTER
- 2 & 3. The development of fine prehension in infancy—B. M. CASTNER
- 3 & 4. The growth of adaptive behavior in infants: An experimental study at seven age levels—H. M. RICHARDSON
- 5 & 6. Differential reactions to taste and temperature stimuli in newborn infants—K. JENSEN

## VOLUME 13—January-June, 1933

1. A critique of sublimation in males: A study of forty superior single men—W. S. TAYLOR
2. A study of the nature, measurement, and determination of hand preference—H. L. KOCH, *et al.*
3. The growth and decline of intelligence: A study of a homogeneous group between the ages of ten and sixty—H. E. JONES AND H. S. CONRAD
4. The relation between the complexity of the habit to be acquired and the form of the learning curve in young children—M. L. MATTHEWSON
5. Eating habits in relation to personality development of two- and three-year-old children: A study of sixty-nine children in two nursery schools—A. A. ELIOT
6. Coordinating mechanisms of the spinal cord—O. C. INGEBRITSEN



# Genetic Psychology: Monographs (continued)

## VOLUME 14—July-December, 1933

1. Mental growth during the first three years: A developmental study of sixty-one children by repeated tests—N. BAYLEY
2. A study of triplets: including theories of their possible genetic relationships—F. N. ANDERSON AND N. V. SCHEIDEMANN
3. The objective measurement of emotional reactions—H. V. GASKILL
4. Development of behavior in the fetal cat—J. D. CONOMOS
5. A study of certain language developments of children in grades four to twelve, inclusive—L. L. LABRANT
6. The effect of early and delayed practice on memory and motor performances studied by the method of co-twin control—J. R. HILGARD

## VOLUME 15—January-June, 1934

1. Studies in the psychology of tone and music—P. R. FARNSWORTH
2. Motor learning of children in equilibrium in relation to nutrition—E. L. BEHR
3. Discrimination limits of pattern and size in the goldfish *Carassius auratus* J. B. ROWLEY
4. Limits of learning ability in the white rat and the guinea pig—B. F. RESS
- 5 & 6. The limits of learning ability in rhesus monkeys—H. A. FJELD

## VOLUME 16—July-December, 1934

1. A statistical study of ratings on the California Behavior Inventory for Nursery-School Children—H. S. CONRAD
2. An eye-movement study of objective examination questions—A. FRANDSEN
3. An experimental study of constitutional types—O. KLINBERG, S. E. ASCH, AND H. BLOCK
4. The development of a battery of objective group tests of manual laterality, with the results of their application to 1300 children—W. N. DUNST
- 5 & 6. An experimental study in the prenatal guinea-pig of the origin and development of reflexes and patterns of behavior in relation to the stimulation of specific receptor areas during the period of active fetal life—L. GANNICHAHL

## VOLUME 17—January-December, 1935

1. Organization of behavior in the albino rat—R. L. THORNDIKE
2. Brightness discrimination in the rhesus monkey—M. F. CRAWFORD
3. The limits of learning ability in cebus monkeys—A. M. KOCH
4. Nature-nurture and intelligence—A. M. LEAHY
5. On intelligence of epileptic children—E. B. SULLIVAN AND L. GARAGAN
6. A study of the play of children of preschool age by an unobserved observer—D. I. COCKRELL

## VOLUME 18—January-December, 1936

1. Sex differences in variational tendency—Q. MCNEAR AND L. M. TERMAN
2. The process of learning to dress among nursery-school children—C. B. KEY, M. R. WHITE, M. P. HONZIK, A. D. HENNEY, AND D. ERWIN
3. A study of the present social status of a group of adults, who, when they were in elementary schools, were classified as mentally deficient—W. R. BALLER
4. The influence of specific experience upon mental organizations—A. ANASTASI
- 5 & 6. Studies in aggressiveness—L. BENDER, S. KEISER, AND P. SCHILDER

## VOLUME 19—January-December, 1937

1. Psychological bases of self-mutilation—C. DABROWSKI
2. Masculine temperament and secondary sex characteristics: A study of the relationship between psychological and physical measures of masculinity—H. GILKINSON
3. A psychological study of forty unmarried mothers—R. D. NOTTINGHAM
4. Behavior problems in the children of psychotic and criminal parents—L. BENDER
5. Domination and integration in the social behavior of young children in an experimental play situation—H. H. ANDERSON
6. The sequential patterning of prone progression in the human infant—L. B. AMES

## VOLUME 20—January-December, 1938

1. The relationship between characteristics of personality and physique in adolescents—P. S. DE Q. CABOT
2. Behavior problems of elementary school children: A descriptive and comparative study—I. Y. MARTEN
3. Graphic representation of a man by four-year-old children in nine prescribed drawing situations—P. F. CHIDLEY
4. Differences between two groups of adult criminals—R. S. TOLMAN
5. A comparative study by means of the Rorschach method of personality development in twenty pairs of identical twins—E. TAUF
6. Individual differences in the facial expressive behavior of preschool children: A study by the time-sampling method—C. SWAN

## VOLUME 21—January-December, 1939

1. An experimental analysis of "level of aspiration"—R. GOULD
2. Some light on the problem of bilingualism as found from a study of the progress in mastery of English among preschool children of non-American ancestry in Hawaii—M. E. SMITH
3. Domination and social integration in the behavior of kindergarten children and teachers—H. H. ANDERSON
4. The capacity of the rhesus and cebus monkey and the gibbon to acquire differential response to complex visual stimuli—W. E. GALT
5. The social-sex development of children—E. H. CAMPBELL

## VOLUME 22—January-December, 1940

1. Measuring human relations: An introduction to the study of the interaction of individuals—E. D. CHAPPEL
2. Aggressive behavior in young children and children's attitudes toward aggression—M. D. FIVE
3. Student attitudes toward religion—E. NELSON
4. The prediction of the outcome-on-furlough of dementia praecox patients—J. S. JACOB
5. Significant characteristics of preschool children as located in the Conrad inventory—K. H. READ
6. Learning by children at noon-meal in a nursery school: Ten "good" eaters and ten "poor" eaters—J. B. McCAY, E. B. WARING, AND P. J. KRUSE
7. Studies in the interpretation of play: I. Clinical observation of play disruption in young children—E. H. ERIKSON

## VOLUME 23—January-June, 1941

1. An analysis of certain variables in a developmental study of language—F. M. YOUNG
2. Infant development under conditions of restricted practice and of minimum social stimulation—W. DENNIS
3. An analysis of the mental factors of various age groups from nine to sixty—B. BALINSKY
4. Factors influencing performance on group and individual tests of intelligence: I. Rate of work—M. W. BENNETT
5. Individual differences in apperceptive reaction: A study of the response of preschool children to pictures—E. W. AMEN

## VOLUME 24—July-December, 1941

1. Twins T and C from infancy to adolescence: A biogenetic study of individual differences by the method of co-twin control—A. GESELL AND H. THOMPSON
2. Finger nail-biting: Its incipency, incidence, and amelioration—A. L. BILLIG
3. An experimental study of the factors of maturation and practice in the behavioral development of the embryo of the frog, *Rana pipiens*—A. FROMME
4. The Fels child behavior scales—T. W. RICHARDS AND M. P. SIMONS
5. Measurement of the size of general English vocabulary through the elementary grades and high school—M. K. SMITH
6. Stereotypes in the field of musical eminence—P. R. FARNSWORTH

## VOLUME 25—January-June, 1942

1. A study of factors determining family size in a selected professional group—J. C. FLANAGAN
2. A genetic study of geometrical-optical illusions—A. WALTERS
3. Interpretation of behavior-ratings in terms of favorable and unfavorable deviations: A study of scores from the Read-Conrad Behavior Inventory—K. H. READ AND H. S. CONRAD
4. Are there any innate behavior tendencies?—J. B. SCHOELLAND
5. An investigation of the intelligibility of the speech of the deaf—C. V. HODGINS AND F. C. NUMBERS

# Genetic Psychology Monographs (continued)

VOLUME 26—July-December, 1942

1. The critical frequency limen for visual flicker in children between the ages of 6 and 18—V. L. MILLER  
Some factors determining handedness in the white rat—K. L. WENTWORTH
2. Motivation and behavior—E. FRENKEL-BRUNSWIK

VOLUME 27—January-June, 1943

1. Comparison of children's personality traits, attitudes, and intelligence with parental occupation—N. R. MADDY
2. A comparative study of mental functioning patterns of problem and non-problem children seven, eight, and nine years of age—M. L. PIGNATELLI

VOLUME 28—July-December, 1943

1. Separation anxiety in young children: A study of hospital cases—H. EDELSTON
2. Correlates of vocational preferences—W. A. BRADLEY, JR.

VOLUME 29—January-June, 1944

1. Mental changes after bilateral prefrontal lobotomy—S. D. PORTUS AND R. D. KEFNER
2. A twin-controlled experiment on the learning of auxiliary languages—B. PRICE, W. J. KOSTIR, AND W. M. TAYLOR

VOLUME 30—July-December, 1944

1. A method of administering and evaluating the thematic appreciation test in group situations—R. M. CLARK
2. A study of anxiety reactions in young children by means of a projective technique—R. TEMPLE AND E. W. AMEN

VOLUME 31—January-June, 1945

1. The evolution of intelligent behavior in rhesus monkeys—B. WEINSTEIN
2. Perceptual behavior of brain-injured, mentally defective children: An experimental study by means of the Rorschach technique—H. WERNER

VOLUME 32—July-December, 1945

1. A clinical study of sentiments: I.—H. A. MURRAY AND C. D. MORGAN
2. A clinical study of sentiments: II.—H. A. MURRAY AND C. D. MORGAN

VOLUME 33—January-June, 1946

1. Interpretation of spontaneous drawings and paintings—T. S. WARENEER  
Preferences for sex symbols and their personality correlates—K. FRANK
2. Outstanding traits in a selected college group, with some reference to career interests and war records—F. L. WELLS AND W. L. WOODS

VOLUME 34—July-December, 1946

1. The relation of emotional adjustment to intellectual function—J. L. DESPERT AND H. O. PIERCE  
The smiling response: A contribution to the ontogenesis of social relations—R. A. SPITZ
2. Finger-painting and personality diagnosis—P. J. NAPOLI

VOLUME 35—January-June, 1947

1. The thematic apperception technique in the study of culture-personality relations—W. E. HENRY
2. A continuation study of anxiety reactions in young children by means of a projective technique—M. DORKEY AND E. W. AMEN

A study of the vocational interest trends of secondary school and college women—A. M. CAWLEY

VOLUME 36—July-December, 1947

1. Maze test validation and psychosurgery—S. D. PORTUS AND H. N. PETERS
2. The diagnostic implications of Rorschach's test in case studies of mental defectives—I. JOLLES

VOLUME 37—January-June, 1948

1. The radio day timer series: A symbolic analysis—W. L. WARNER AND W. E. HENRY  
The relation of personality characteristics and response to verbal approval in a learning task—G. L. GRACE
2. The mechanism of vision: XVIII. Effects of destroying the visual "associative areas" of the monkey—K. S. LASHLEY  
A study of the relationship between handwriting and personality variables—P. CASTELNUOVA-TEDESCHI

VOLUME 38—July-December, 1948

1. Modern language learning: The intensive course as sponsored by the United States Army, and implications for the undergraduate course of study—M. LINN
2. Schizophrenia and the MAPS test: A study of certain formal psycho-social aspects of fantasy production in schizophrenia as revealed by performance on the Make a Picture Story (MAPS) Test—E. S. SHNEIDMAN  
A study of the transmission of authority patterns in the family—H. L. INGRAMSOLL

VOLUME 39—January-June, 1949

1. A study of the psychoanalytic theory of psychosexual development—G. S. BLUM  
The assessment of parental attitudes in relation to child adjustment—E. J. SHOEN, JR.
2. Qualitative differences in the vocabulary responses of normals and abnormal—H. FRIFEL  
The relative effectiveness of motion and still pictures as stimuli for eliciting fantasy stories about adolescent-parent relationships—P. E. EISENER  
The organization of hereditary maze-brightness and maze-dullness—L. V. SEARLE

VOLUME 40—July-December, 1949

1. An experimental study of what young school children expect from their teachers—B. BIBER AND C. LEWIS  
A study of the relative effects of age and of test difficulty upon factor patterns—H. A. CURTIS
2. A projective experiment using incomplete stories with multiple choice endings—J. K. SEATON  
Effects of sex role and social status on the early adolescent personality—E. MILNER  
Social perceptions and attitudes of children—M. RADKE, H. TRACER, AND H. DAVIS

VOLUME 41—January-June, 1950

1. Some psychological and educational aspects of pediatric practice: A study of well-baby clinics—L. H. BLUM  
One-trial learning in the domestic rat—B. B. HUDSON  
An introduction to the principles of scientific psychoanalysis—A. ELLIS
2. Awareness of racial differences by preschool children in Hawaii—D. V. SPRINGER  
Age trends in children's evaluation of teacher-approved and teacher-disapproved behavior—S. L. WITRYOL  
The relationship between level of vocational aspiration and certain personal data: A study of some traits and influences bearing on the prestige level of vocational choice—J. STUBINS

VOLUME 42—July-December, 1950

1. Personality patterns of suicidal mental hospital patients—N. L. FARBROW  
Sex-role identification in young children in two diverse social groups—M. RABBAN
2. A study of the influence of the social field on individual behavior: As revealed in the expression of hostility and warmth by neurotic and paranoid schizophrenics in discussion group situations—D. SHAPIRO  
An experimental study of avoidance—R. F. HEFFERLINE

VOLUME 43—January-June, 1951

1. A study of copying ability in children—E. A. TOWNSEND  
Prestige motivation of gifted children—D. P. AUSUBEL
2. A psychological study of physical scientists—A. ROE

VOLUME 44—July-December, 1951

1. The organization of hostility controls in various personality structures—S. FISHER AND E. HINDS  
Children and radio: A study of listeners and non-listeners to various types of radio programs in terms of selected ability, attitude, and behavior measures—E. A. RICCIUTI
2. Quantitative expression in young children—W. E. MARTIN  
The use of magnetic devices in the collection and analysis of the preverbal utterances of an infant—A. W. LYNF

VOLUME 45—January-June, 1952

1. Japanese-American personality and acculturation—W. CAUDILL
2. A statistical study of the Freudian theory of levels of psychosexual development—C. A. BARNES  
Personality characteristics of selected disability groups—D. N. WIEN

# Genetic Psychology Monographs (continued)

VOLUME 46—July-December, 1952

1. The relationship of social status, intelligence, and sex of ten- and eleven-year-old children to an awareness of poverty—F. J. ESTYAN  
An empirical study of the castration and Oedipus complexes—S. M. FRIEDMAN
2. The relationship between projective test scoring categories and activity preferences—M. M. SCHWARTZ  
A comparison of formal and content factors in the diagnostic testing of schizophrenia—M. SHERMAN

VOLUME 47—January-June, 1953

1. Ability and accomplishment of persons earlier judged mentally deficient—D. C. CHARLES  
Variations in the consistency of the behavioral meaning of personality test scores—M. KORNREICH
2. Some child-rearing antecedents of aggression and dependency in young children—R. R. SEARS, *et al.*  
Symptom correlates for descriptive diagnosis—J. R. WITTENBORN, *et al.*

VOLUME 48—July-December, 1953

1. Age and mental abilities: A longitudinal study—W. A. OWENS, JR.  
The development of a personality questionnaire for drinkers—P. J. HAMPTON  
Personality and physical disease: A test of the Dunbar hypothesis applied to diabetes mellitus and rheumatic fever—D. H. CROWELL
2. Socio-economic contrasts in children's peer culture prestige values—B. POPK  
A critical review of the stability of social acceptability scores obtained with the partial-rank-order and the paired-comparison scales—S. A. WITRYOL AND C. G. THOMPSON  
A study of the effects of color on Rorschach responses—G. C. BRODY

VOLUME 49—January-June, 1954

1. Factors underlying major reading disabilities at the college level—J. A. HOLMES  
Parent behavior toward first and second children—J. K. LASKO
2. Social-status and intelligence: An experimental study of certain cultural determinants of measured intelligence—E. A. HAGGARD  
Certain determinants and correlates of authoritarianism—S. SIEGEL  
Personalities in faces: I. An experiment in social perceiving—P. F. SECORD, W. F. DUKES, AND W. BEVAN

VOLUME 50—July-December, 1954

1. A study of the relationship between play patterns and anxiety in young children—E. W. AMEN AND N. RENISON  
Operational exploration of the conceptual self system and of the interaction between frames of reference—M. EDELSON AND A. E. JONES  
Problem solving: A statistical description of some relationships between organismic factors and selected response measures—N. A. FATTU, E. KAPOR, AND E. V. MECH
2. The relation of cortical potentials to perceptual functions—C. CHYATTE  
The impact for clinical psychology of the use of tests derived from theories about infantile sexuality and adult character—D. W. MILES  
Measuring personality in developmental terms: The Personal Preference Scale—M. H. KROUT AND J. K. TABIN

VOLUME 51—January-June, 1955

1. Some relations between techniques of feeding and training during infancy and certain behavior in childhood—A. BERNSTEIN  
The expression of personality in drawings and paintings—L. H. STEWART
2. Negative stereotypes concerning Americans among American-born children receiving various types of minority-group education—J. A. FISHMAN  
The Lincoln-Oseretsky Motor Development Scale—W. SLOAN

VOLUME 52—July-December, 1955

1. Some personality correlates of sex, sibling position, and sex of sibling among five- and six-year-old children—H. L. KOCH  
A quantitative Rorschach assessment of maladjustment and rigidity in acculturating Japanese Americans—G. DE VOS  
Measurement of authoritarianism and its relation to teachers' classroom behavior—H. M. MCGEE
2. The formal aspects of schizophrenic verbal communication—B. MIRIN  
A study in an aspect of concept formation, with subnormal, average, and superior adolescents—H. N. HOFFMAN  
Traumatic avoidance learning: Acquisition and extinction in dogs deprived of normal peripheral autonomic function—L. C. WYNN AND R. L. SOLOMON

VOLUME 53—January-June, 1956

1. As the psychiatric aide sees his work and problems—F. L. WELLS, M. GREENBLATT, AND R. W. HYDE  
An investigation of avoidance, anxiety, and escape behavior in human subjects as measured by action potentials in muscle—J. D. BROTHERS  
Spread of effect: A critical review—M. H. MARX
2. Stress, fantasy, and schizophrenia: A study of the adaptive processes—O. J. B. KERNER  
The attitude structure of the individual: A Q-study of the educational attitudes of professors and laymen—F. N. KERLINGER

VOLUME 54—July-December, 1956

1. A study of personality differences between middle and lower class adolescents: The Szondi Test in culture-personality research—L. RAINWATER  
The assessment of parental identification—S. W. GRAY AND R. KLAUS
2. The influence of social context on impulsive and control tendencies in preadolescents—G. H. ZUK  
Tender-mindedness versus tough-mindedness in psychology: A reexamination—H. WINTHROP  
A method for the comparison of groups: A study in thematic apperception—L. C. SCHAW AND W. E. HENRY

VOLUME 55—January-June, 1957

1. Academic performance and personality adjustments of highly intelligent college students—B. M. HORRALL  
The use of the Vineland Social Maturity Scale in the planning of an educational program for non-institutionalized low-grade mentally deficient children—M. G. GOTTSSEN
2. The structure and origin of the anal character—H. BELOFF  
Free expression of adolescents' interests—M. AMATORA  
The role of mass media and the effect of aggressive film content upon children's aggressive responses and identification choices—R. S. ALBERT  
Interest in persons as an aspect of sex difference in the early years—E. W. GOODENOUGH

VOLUME 56—July-December, 1957

1. Some stable response determinants of perception, thinking, and learning: A study based on the analysis of a single test—M. L. SIMMEL AND S. COUNTS  
Potentials of age: An exploratory field study—S. L. PRESSEY
2. Tactile communication—L. K. FRANK  
The use of a filmed puppet show as a group projective technique for children—M. R. HAWORTH  
The social competence of middle-aged people—R. J. HAVIGHURST

VOLUME 57—January-June, 1958

1. Psychological and cultural problems in mental subnormality: A review of research—S. B. SARASON AND T. GLADWIN
2. Developmental aspects of discrimination in relation to adjustment—P. LONDON  
Muscular tension as a response to psychological stress in rheumatoid arthritis and peptic ulcer—J. A. SOUTHWORTH  
Frustration: Theory and experiment—R. LAWSON AND M. H. MARX

VOLUME 58—July-December, 1958

1. Emotional aspects of political behavior: The woman voter—E. M. BENNETT AND H. M. GOODWIN  
The accuracy of self estimate: A measure of self-concept reality—R. M. BRANDT  
Personality factors in social mobility: A study of occupationally mobile businessmen—J. C. ABREGLEN
2. Interaction between authoritarian and nonauthoritarian principals and teachers—P. LAMBERT  
Some factors related to the choice-status of ninety eighth-grade children in a school society—D. ELKINS  
Some relationships between fantasy and self-report behavior—R. C. CALOGERAS



# Genetic Psychology Monographs (continued)

VOLUME 59—January-June, 1959

1. Freshman Marchbooks and college performance: G. R. SCHWENGER, M. J. NELSON, and M. BRISTOL  
Perceptual aspects of attributed movement: R. W. COAN  
Men and women: Personality patterns and contrasts: J. M. BENNETT and L. R. COHEN  
2. Personality factors in mothers of cerebral palsy children: C. BOKS  
Bipartism and integration: A social psychological analysis of editorial content in New York newspapers of the  
American authority groups: A. FISHMAN and G. S. FISHMAN  
Self, role, and satisfaction: A. L. BOORST

VOLUME 60—July-December, 1959

1. The constancy of personality ratings over two decades: H. D. THORNBAM  
Motherhood, identification, and emotional interest: I. H. STEINMAN  
A study of mother-child relationships in the emotional disorders of children: M. J. ROSENTHAL, M. FINKELSTEIN,  
N. and R. I. ROSENTHAL  
On the trail of the mother-child: W. F. OLDS and N. K. ROSE  
2. Measuring the mental health of normal adults: R. J. FELD  
Parental identification in young children: W. FINKELSTEIN  
The problems of Jewish religious youth: B. M. LAYMAN

VOLUME 61—January-June, 1960

1. Identifying socially maladjusted school children: A. A. ADAMS  
Social interaction in an institutional nursery school: H. W. STEVENSON and N. G. STEVENSON, et al  
Interests of pre-adolescent boys and girls: M. AUSTON  
Educational and sex differences in the organization of ability in technical and academic students in Colombia, South  
America: J. P. VILLAS  
2. Some changes in the personalities of German men: L. HAINSWAY  
Right-left gradients in body image, body consistency, and perception: S. FISHER  
Longitudinal studies of child research responses: Younger subjects two to 10 years: L. B. AMES  
Personality patterns related to occupational roles: M. SIMONSON and H. F. PUCK

VOLUME 62—July-December, 1960

1. The reading process: A descriptive interdisciplinary approach: M. P. LANGMAN  
Responses to stress as a function of the level of intelligence: S. M. H. ZAINI  
Research design: A systematic combining of content, thought process and determinant scales: P. A. BOWEN  
N. TOSHI and A. KAWADA  
Longitudinal studies of child research responses: Older subjects aged 10 to 16 years: J. B. AMES  
2. Twenty years of child therapy in America, 1937-1956: An annotated bibliography: N. H. FINKELSTEIN, R. NITTELBY and  
A. BORN  
The quantitative analysis of parent behavior toward psychotic children and their siblings: F. M. DONNELLY

VOLUME 63—January-June, 1961

1. Make a conscious trait: An approach to objective scoring of sentence completions: J. P. BOWEN  
Some aspects of the relationship between perception and motility in children: E. P. ROTHMAN  
Some psychological aspects of mental illness in children: P. H. BOWEN  
2. Some material changes in children's personality and character: J. C. FISHER  
A study of change formation in juvenile mental defective and brain damaged adults: H. W. ZASLOW  
An investigation of trait relationships among six year old children: S. A. AVERIAN

VOLUME 64—July-December, 1961

1. Coping patterns of pre-hand children in response to intelligence test demands: A. MUMFORD  
The role of individual and paired interactional problem solving behavior of rats: H. SOLITARY and social animals  
A. BORN and B. A. FINKELSTEIN  
2. Studies of role performance: I. H. MANN  
A psychological assessment of psychomotor and related performance: R. TAPP  
A cross-sectional study of emotional status: W. N. STEINMAN  
The biological basis of creativity: R. GUTMAN

VOLUME 65—January-June, 1962

1. A research study of the development of personality structure in white and Negro children in a Southwestern  
community: A. I. FINE  
Measuring personality: Patterns of emotion: I. FINKELSTEIN  
Some interests in new children: M. FINKELSTEIN  
Social and emotional adjustment during adolescence as related to the development of psychomotoric illness in  
adolescence: I. H. MANN  
2. A cross-sectional study of interaction process: Stress, peer and self assessments, personality and other variables: E. F.  
CHILDREN'S CONCEPTIONS OF THE CAUSES AND FUNCTIONS OF THE HUMAN BODY: J. GILBERT

VOLUME 66—July-December, 1962

1. Greater manual dexterity and perceptual reaction of lateral criminals: L. FINKELSTEIN  
Perceptual ability among children and their reactions to children's hostility: A. R. FINKELSTEIN  
Some explanations of the nature of children's hostility: A. R. FINKELSTEIN  
2. Developmental characteristics of responses to emotion: Some illustrations: J. R. FINKELSTEIN and I. H. MANN  
Thinking as a function of social: An experiment in early childhood learning: W. FINKELSTEIN  
Achievement goals in science during the elementary years and over the summer vacation period: A study of trends in  
achievement by sex and grade level among students of average intelligence: K. M. FINKELSTEIN, Jr. and M. FINKELSTEIN

VOLUME 67—January-June, 1963

1. Verbal abilities of a mother and child in therapy: E. FINKELSTEIN  
A study of intelligence: Differences from primary to selected self-concept: W. I. FINE  
On integrating study of the human structure of adult intelligence: R. N. CATTALL and J. HENRY  
2. Conceptual performance associated with degree of quantitative and non-quantitative college majors: J. GILBERT  
A study of the concept of the human role of 11 middle-class American families: A. STEINMAN

VOLUME 68—July-December, 1963

1. Main dimensions as reflected in intelligence: G. A. YERKES  
The nature and measurement of components of motivation: R. B. CATTALL, J. A. HANCOCK, and A. S. SWENST  
2. Changes in preschool children as a function of content and type of reinforcement: W. F. FISHER  
The Cornell Experiment: New York as a research developmental study: I. R. AMES and J. I. ILL  
Psychological and sociological factors in prediction of group patterns of women: M. C. MILBY  
Functionalism and its historical age & cause: R. HANCOCK





## PUBLICATION IN THE JOURNAL OF PSYCHOLOGY

*The Journal of Psychology* provides a special service to authors and readers. Articles accepted for publication appear in print with the least possible delay. Although editorial standards are the same as those for the other journals, all correspondence and handling are expedited, and the referees may, at their discretion, permit some latitude in length of treatment with respect to tabular material, discussion, and review of the literature.

To save time, authors are requested to exercise unusual care in preparing their papers. Special attention should be paid to the General Instructions on the back cover of each issue and to the announcement on the first preliminary page.

\$39.00 per annum  
\$18.00 per volume  
Single Numbers \$9.00

BIMONTHLY  
Three volumes per year  
Immediate publication

September, 1973  
Volume 85, First Half

(Founded in 1935 by Carl Murchison)

# The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a bilateral contract, and may not be resold for a period of five years from the date of publication.

SEPTEMBER, 1973

(Manuscripts are printed in the order of final acceptance)

An experimental test of the contingency model on small classroom groups . . . . .	3
BY ROBERT C. HARDY, STANLEY SACK, AND FRANCES HARPINE	
Is test anxiety rising? . . . . .	17
BY ARLINE L. BRONZAFT AND GILDA F. EPSTEIN	
The astrological "theory" of personality: An unbiased test by a biased observer . . . . .	21
BY ROBERT J. PELLEGRINI	
Failure to find relationships between family constellation and achievement motivation . . . . .	29
BY D. J. W. STRÜMPFER	
Self-actualization and self-ideal discrepancy . . . . .	37
BY JOHN MAHONEY AND JACK HARTNETT	

(OVER)

Copyright © 1973 by The Journal Press  
Provincetown, Massachusetts, U. S. A.

All rights reserved

Second-class postage paid at Provincetown, Mass  
and additional mailing offices

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

The control orientation of black athletes in relation to unobtrusive tasks of skill . . . . .	43
By STEVEN THURBER, DEL HEACOCK, AND BOOKER BROWN	
The relationship of marijuana usage to personality and motivational factors . . . . .	45
By CHARLES BURDSAL, GARY GREENBERG, AND RANDIE TIMPE	
Humor judgments as a function of identification classes: Canadian vs. American . . . . .	53
By LAWRENCE LA FAVE, KEVIN MCCARTHY, AND JAY HADDAD	
Birth order and related variables in a large outpatient population . . . .	61
By BERNARD F. RIESS AND JEANNE SAFER	
Implicit time constraints in the measurement of productive thinking . . .	69
By ROBERT J. MENGES AND BELINDA JOHNSON	
Birth order, personality development, and vocational choice of becoming a Carmelite nun . . . . .	75
By PHILIP S. VERY, ROBERT B. GOLDBLATT, AND VINCENT MONACELLI	
Development of a psychological effectiveness scale . . . . .	81
By CHARLES A. POB	
Some international comparisons of cancer mortality rates and personality: A brief note . . . . .	87
By GORDON RAE AND JAMES MCCALL	
Randomization tests: Computer time requirements . . . . .	89
By EUGENE S. EDGINGTON AND ALLAN R. STRAIN	
Catholic sisters and the Edwards Personal Preference Schedule . . . .	97
By HARRY W. GARDINER	
Personality and demographic correlates of simulated personal space . .	101
By DARHL M. PEDERSEN	
Semantic style variance in personality questionnaires . . . . .	109
By ROBERT STRAHAN AND KATHLEEN CARRESE GERBASI	
Relations among physical attractiveness, body attitudes, and self-concept in male and female college students . . . . .	119
By RICHARD M. LERNER, STUART A. KARABENICK, AND JOYCE L. STUART	
In defense of measuring attitudes . . . . .	131
By MARION STEININGER	
A device for investigating adaptation to sensory rearrangement . . . .	137
By JAMES R. LACKNER	
Modeling and personal space behavior in children . . . . .	143
By KENT G. BAILEY, JOHN J. HARTNETT, AND HILDA W. GLOVER	
Personality correlates of proximity preferences . . . . .	151
By ALAN F. SEWELL AND JAMES T. HEISLER	
Person perception as a function of the personal consequences and immediacy of a decision . . . . .	157
By ALAN F. SEWELL	
The EPSAT as a predictor . . . . .	165
By JERRY W. O'DELL	

## AN EXPERIMENTAL TEST OF THE CONTINGENCY MODEL ON SMALL CLASSROOM GROUPS\*

*University of Maryland and Catholic University*

---

ROBERT C. HARDY, STANLEY SACK, AND FRANCES HARPINE<sup>1</sup>

---

### SUMMARY

The problem was to test the applicability of Fiedler's contingency model on 56 junior high school classroom groups. It investigated the effectiveness of high and low least preferred co-worker (LPC) leaders with good and poor leader member relations (LMR) and weak power on structured and unstructured group tasks. The results indicated that high LPC leaders were more effective for a structured task when LMR were poor, and no relationship was found when LMR were good. For an unstructured task the results indicated that high LPC leaders were more effective when LMR were good. When LMR were poor, low LPC leaders were more effective than high on only one of the four measures used. These findings offer partial support for the model.

---

### A. INTRODUCTION

One of the most controversial theories of leadership effectiveness is Fiedler's (5) contingency model. To measure leadership style Fiedler used the Least Preferred Co-worker (LPC) scale as cited in Fiedler (7). This scale is a 16-20 item bipolar adjective checklist in which individuals rate a co-worker with whom they have worked least well. The lower they score on this scale the more they associate poor task success with the characteristics of the person they are rating, and the more task-oriented is their leadership style. The higher they score the more they disassociate task failure from the personality characteristics of the person they are rating, and the more relation-oriented is their leadership style. High LPC leaders, when faced with a choice, seek close relationships rather than concentrating on the task. If the situation is favorable to them, and if these relationships are not threatened, these leaders concentrate on their secondary goal of task success. Low LPC leaders when faced with a choice concentrate on task success over personal relationships.

---

\* Received in the Editorial Office on May 8, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Direct requests for reprints to the first author at the address shown at the end of this article. The third author is at Catholic University.

If they are in a favorable situation and task success is not threatened, they concentrate on improving their secondary goal of the relationship within their group (9).

Fiedler (5) stated that leadership effectiveness was a function of situation favorability, where favorability referred to the leader's ability to influence and control the group. He maintained that the favorability of the situation was based on three variables which, in the order of importance, were leader member relations (LMR), task structure, and position power. By dichotomizing each of these variables he arrived at eight possible cells or situations ranging from most favorable to least favorable. The above can best be understood by examining Table 1. In this table Cell 1 represents the most favorable situation, while Cell 8 represents the least favorable situation.

TABLE 1  
CONTINGENCY MODEL

Cell	Leader member relations	Task structure	Position power
1	Good	Structured	Strong
2	Good	Structured	Weak
3	Good	Unstructured	Strong
4	Good	Unstructured	Weak
5	Poor	Structured	Strong
6	Poor	Structured	Weak
7	Poor	Unstructured	Strong
8	Poor	Unstructured	Weak

By combining the results of a large number of studies Fiedler (7) postulated that a U-shaped function existed between leadership effectiveness, as measured by group success, and task orientation of the leader, as measured by the Least Preferred Co-worker (LPC) scale. Specifically, low LPC leaders were more effective in Cells 1, 2, 3, and 8, while high LPC leaders were more effective in Cells 4 and 5. Cell 6 had not been specifically investigated, and the results concerning Cell 7 were inconsistent and did not strongly support either high or low LPC leaders. Thus, in extremely favorable or unfavorable situations low LPC leaders were more effective than high, while in moderately favorable situations high LPC leaders were more effective than low.

Tests of the contingency model were made by Anderson (1); Shaw and Blum (18); Hunt (13); Fiedler (6); Reilly (16); Hill (12); O'Brien, Fiedler, and Hewett (15); Graen, Alvares, Orris, and Martella (10); Hardy (11); and Chemers and Skrzypek (3). The studies by Anderson and Shaw and Blum offered only weak support for the model, while those by Reilly and Graen *et al.* failed to replicate any parts of the model. The remaining studies offered strong support for the model.



Fiedler (8) criticized the methodology in the Graen *et al.* (10) study, stating that the investigators improperly manipulated situation favorability. In particular, they failed to differentiate task structure, the power of the leader was extremely weak, and leader member relations differed considerably across conditions. The Chemers and Skrzypek (3) study was designed to investigate all eight cells of the model and was prompted by the controversy surrounding the Graen *et al.* article. Particular emphasis was given to the careful manipulation of the situation dimensions which were criticized by Fiedler in the previously discussed study. Their results strongly supported the model; however, the correlation for Cell 3 was in the reverse direction, and the correlation for Cell 6 was only  $+13$ .

The Hardy study (11) was designed to investigate the first four cells of the model. His results supported the model, except that the results for Cell 2 were in the reverse direction. These results were consistent with five out of seven laboratory experiments and inconsistent with three out of three field experiments reviewed in Mitchell, Beglen, Onchen, and Fiedler (14). Two possible explanations were proposed in the Hardy study. First, that the degree of structure of the structured task was not differentiated from the unstructured task. Second, that position power was a more decisive determinant of situation favorableness than was task structure. If this were the case, Cells 2 and 3 would reverse positions in the hierarchy, as well as Cells 6 and 7.

The present study was designed specifically to investigate Cell 2, where results have been so inconsistent, and its complement with poor relations, Cell 6. In addition to these two cells, Cells 4 and 8, which differed only in respect to task structure, were also investigated. The entire model was not investigated, since a larger sample was not available. To reduce the size of the sample per cell in this study would have increased the probability of making a Type II error to a point that was not deemed satisfactory, especially in view of the argument concerning sample sizes in replication experiments made by Tversky and Kahneman (21) and discussed in detail by Shiflett (19).

A weakness in many of the studies investigating the contingency model was the reliance on correlations as the mode of analysis. Shiflett (19) demonstrated, using data from a variety of studies concerning the contingency model, that when analysis of variance was used, the precision of the experiments was increased. Further, of the 72 correlations Fiedler (5) used to support his theory only 15 were significant at the .05 level or less (2). Consequently, the use of correlations as a mode of analysis was avoided.

On the basis of the weak power cells of the contingency model the hypotheses investigated were as follows:

1. An interaction effect exists between LPC and LMR when the task is structured. (1a) Low LPC leaders are more effective than high LPC leaders when the task is structured and the LMR are good. (1b) High LPC leaders are more effective than low LPC leaders when the task is structured and the LMR are poor.

2. An interaction effect exists between LPC and LMR when the task is unstructured. (2a) High LPC leaders are more effective than low LPC leaders when the task is unstructured and the LMR are good. (2b) Low LPC leaders are more effective than high LPC leaders when the task is unstructured and the LMR are poor.

## B. METHOD

### 1. *Subjects*

Subjects were 224 (95 males and 129 females) students who were members of seven sections of ninth grade social studies at an Eastern suburban junior high school.

### 2. *Procedure*

One to two weeks prior to the experimental session, during a regularly scheduled class period, all subjects were told by their teacher that in the near future they were going to do some group work. As an aid to organizing the groups, all subjects were asked to complete the LPC scale and a sociometric scale in which they were asked to list the four classmates with whom they would most like to work, and the four with whom they would least like to work on the group assignment. In each section four high and four low LPC-led groups were formed, with one leader and three group members each. Two of the high and two of the low LPC-led groups had good leader member relations, while the rest had poor relations. In all groups the leader with the most extreme score possible with the appropriate LMR was selected. There were a total of 14 high and 14 low LPC-led groups with good LMR, and the same with poor LMR. Thus, there were two replications within each section and 14 replications in total.

To determine if the actual groups formed had the appropriate LMR, the leader completed the group atmosphere scale, and the group members completed the group evaluation scale after the last task. The group evaluation scale measures the group member's perception of his ability to work with the leader, his desire to remain in the group, his belief concerning how the rest of the group members perceive the effectiveness of the leader, and his perception concerning the extent to which he is an integral part of the group. The

group atmosphere scale measures the leader's perception of the overall feelings of the group members toward the group.

In all groups, leaders were given a weak power communication. The teacher informed all students that the leader was picked for no particular reason over any other group member, and that for future work with the group the group members could select another group leader. To validate this manipulation the position power scale was administered to all group members except the leader at the conclusion of the last task.

All groups were assigned to both a structured and an unstructured task concerned with a 12 minute, open ended, stimulus film, *What Do I Know About Benny?* (22). Task structure was defined by decision verifiability, goal clarity, goal-path multiplicity, and solution multiplicity, as cited in Shaw (17). A structured task was high in decision verifiability and goal clarity, and low in solution multiplicity and goal-path multiplicity. The unstructured task had the opposite characteristics. The order of presentation of the tasks was counterbalanced to control for an order effect. Groups were randomly assigned to do either the structured or the unstructured task the first day, and the second day they were given the opposite task. The time limit for each was 35 minutes. The score on these tasks was the dependent variable used in this study. At all times during the course of the experiment a genuine classroom atmosphere was maintained in which the students were led to believe that the work they were completing was part of their regular assignments and would contribute to their course grades.

The unstructured task was an opinion-type question concerning recommendations the group (which was to assume it was a student committee concerned with student affairs) was to make to the principal concerning Benny, a student discussed in the stimulus film. The results were evaluated by three judges on the basis of three criteria: style, adequacy of recommendations, and persuasibility. Style referred to the grammatical correctness of the paper, the correctness of the spelling, and the appropriateness of the language used. Adequacy of recommendations referred to the degree to which the recommendations adequately covered all aspects of the problem. Persuasibility referred to the degree to which the principal would be influenced to accept the recommendations presented in the paper.

The structured task was an objective-type questionnaire dealing with details and problems presented in the stimulus film. The group was instructed to work on only one question at a time. They were to reach group consensus before marking the answer on an answer sheet. It was evaluated by summing the number of correct responses to the test.

### 3. Design

In each section the teacher informed the students that they were about to do some group work. He then told them they were about to look at the film *What Do I Know About Benny?*, which is concerned with a parent-teacher conference. After the film was completed, the teacher formed the groups and appointed the group leaders. Students not needed for the experiment were placed into additional groups; however, their results were not included in the analysis. At this time the weak power communication was given to all students, and they were informed that since the assignments they were about to do would count toward their final grades, they should do their best. The first task was then given to the group leader. At the end of the time limit all questionnaires were collected, and the groups were informed they would do another assignment during their next class period, which was one day later. During the second day of the experiment, the teacher, after the groups were organized, handed out the second task to the group leaders. When this task was completed and collected, the leader was given the group atmosphere scale to complete, while the group members were given the group evaluation scale and the position power scale to complete.

The unstructured task was retyped exactly as it had been written. Copies of the task were randomly arranged for each of the judges, who were advanced graduate students. The judges received a standard instruction sheet which defined the three criteria by which the task was to be scored and the procedures by which the scoring was to be done. They were to evaluate the papers on each criterion using any value from one to 50, with one representing a poor evaluation and 50 a good one. Further, all papers were to be graded on a single criterion before any paper was to be graded on the next criterion. To evaluate the structured task the groups' results were scored by an objective key.

### 4. Analysis of Data

To analyze the data a  $2 \times 2$  factorial design between LPC (high and low) and leader member relations (good and poor) was used for both the structured and the unstructured task. Hypothesis 1, where the dependent variable was the groups' effectiveness on the structured task, tested Cells 2 and 6 of the contingency model. An interaction effect was predicted in which the low LPC leaders would be more effective than the high when the relations were good, and the high LPC leaders would be more effective than the low when the relations were poor. Hypothesis 2, where the dependent variable was the groups' effectiveness on the unstructured task, investigated Cells 4 and 8



of the contingency model. An interaction effect was predicted, in which the high LPC leaders would be more effective than the low when the relations were good, and the low would be more effective than the high when the relations were poor. To test hypotheses 1a, 1b, 2a, and 2b, which investigated Cells 2, 6, 4, and 8, respectively, an individual comparison was run between the appropriate cells by use of the  $F$  test cited in Winer (23, p. 120).

### C. RESULTS

To determine if the groups used in this study had the appropriate LMR, two scales were administered at the conclusion of the experiment. The first of these, the group atmosphere scale, represented the leader's perception of his relationship with the group. The possible scores on this scale range from 10 to 80 with the higher value representing the more favorable relations. The mean and standard deviation for the good relations groups were 73.6 and 10.2, respectively, and for the poor relations groups they were 59.9 and 16.2, respectively. A one way analysis of variance was run between the two groups, yielding an  $F$  value ( $df = 1, 54$ ) of 14.2095 ( $p < .001$ ). The second of these scales, the group evaluation scale, represented the group's evaluation of the atmosphere within the group and of the relationship between the group members and the leader. The range of scores on this scale is from 4 to 20, with 4 representing the least favorable relations and 20 the most favorable. To obtain a group score the mean of the three group members (the leader did not complete this scale) was computed. The mean and standard deviation for the groups with good LMR were 16.5 and 2.5, respectively, while they were 12.7 and 3.0, respectively, for the poor LMR groups. A one way analysis of variance between the two groups yielded an  $F$  value ( $df = 1, 54$ ) of 25.7531 ( $p < .001$ ). Thus, the results of the analysis of variance indicated that the two groups were significantly different from each other in respect to LMR. The poor LMR groups did score higher on the scales than was anticipated, and these groups might better be described as moderately poor instead of poor, which is consistent with Fiedler (7). On the basis of the sociometric, group atmosphere and group evaluation data it was concluded that the relationships within the groups were appropriately designated.

To determine if the weak power manipulation was successful, the position power checklist was administered at the conclusion of the experimental session to all group members, excluding the leader. The range of possible scores on this scale is from 5 to 18, with 5 indicating that group members perceive the leader to have very little power. The mean and standard deviation for all completing the scale was 7.9 and 3.0, respectively. This is indicative of rea-



sonable weak position power; thus it was concluded that the power manipulation was successful.

To test hypothesis 1, a  $2 \times 2$  factorial analysis of variance was run between LPC and LMR with the structured task as the dependent variable with 14 cases per cell. The means of these data are presented in Table 2, and the results of the analysis of variance are presented in Table 3. An interaction effect was predicted with low LPC leaders being more effective when the relations were good, and high LPC leaders being more effective when they were poor. A main effect, however, was found with high LPC leaders being more effective than low LPC leaders. Thus, hypothesis 1 was not supported by the data.

TABLE 2  
STRUCTURED AND UNSTRUCTURED TASK MEANS OF HIGH AND LOW LPC  
GROUPS WITH GOOD AND POOR LMR

LPC	Task	LMR	
		Good	Poor
High	Structured	21.357	20.929
	Unstructured		
	Style	52.609	46.591
	Persuasibility	53.334	48.068
	Adequacy of recommendation	51.567	48.842
	Total	52.994	47.143
Low	Structured	18.786	16.286
	Unstructured		
	Style	48.151	52.351
	Persuasibility	47.179	50.742
	Adequacy of recommendation	47.631	48.759
	Total	47.191	51.021

Note: LPC = least preferred co-worker; LMR = leader member relations.

To test hypothesis 1a, which investigated Cell 2 of the contingency model, an individual comparison was made on the structured task between high and low LPC when the LMR were good. It was hypothesized that low LPC leaders are more effective than high LPC leaders. The results of this analysis are presented in Table 4. The  $F$  test ( $df = 1, 52$ ) was not significant at the .05 level. Further, Table 2 shows that the means were in the opposite direction from what was hypothesized. Thus, hypothesis 1a was not supported by the data.

Hypothesis 1b investigated Cell 6 of the contingency model in which high

TABLE 3  
SUMMARY OF ANALYSIS OF VARIANCE FOR FACTORIAL DESIGN FOR HIGH AND  
LOW LPC GROUPS WITH GOOD AND POOR LMR ON THE STRUCTURED  
AND UNSTRUCTURED TASKS

Task	Source of variation	Sum of squares	df	Mean square	F
Structured	LPC	182.16071	1	182.16071	5.276*
	LMR	30.01786	1	30.01786	
	LPC $\times$ LMR	15.01789	1	15.01789	
	Within cell	1795.35681	52	34.52609	
	Total	2022.55327	55		
Unstructured Style	LPC	11.55702	1	11.55702	5.121*
	LMR	5.92800	1	5.92800	
	LPC $\times$ LMR	365.36514	1	365.36514	
	Within cell	3710.00372	52	71.34622	
	Total	4092.85388	55		
Unstructured Persuasi- bility	LPC	10.14904	1	10.14904	3.646*
	LMR	42.38639	1	42.38639	
	LPC $\times$ LMR	272.80289	1	272.80289	
	Within cell	3890.54568	52	74.81819	
	Total	4215.88397	55		
Unstructured Adequacy of recom- mendation	LPC	8.93601	1	8.93601	.968
	LMR	56.54133	1	56.54133	
	LPC $\times$ LMR	51.93648	1	51.93648	
	Within cell	2788.79962	52	53.63076	
	Total	2906.21344	55		
Unstructured Total	LPC	14.31174	1	14.31174	4.475*
	LMR	12.96970	1	12.96970	
	LPC $\times$ LMR	328.00683	1	328.00683	
	Within cell	3811.32669	52	73.29474	
	Total	4166.61493	55		

Note: LPC = least preferred co-worker; LMR = leader member relations.

\*  $p < .05$ .

LPC leaders were hypothesized to be more effective than low LPC leaders. To test this an individual comparison was made between the appropriate means. The results of this analysis are presented in Table 4. The  $F$  test ( $df = 1, 52$ ) was significant, and the means, which are presented in Table 2, were in

TABLE 4  
INDIVIDUAL COMPARISONS BETWEEN HIGH AND LOW LPC LEADERS WITH  
GOOD AND POOR LMR ON STRUCTURED AND UNSTRUCTURED TASKS

Task	LMR	F test
Structured		
	Good	2.681
	Poor	8.741**
Unstructured Style		
	Good	3.900
	Poor	6.509*
Unstructured Persuasibility		
	Good	7.087*
	Poor	1.339
Unstructured Adequacy of Recommendation		
	Good	4.044*
	Poor	.002
Unstructured Total		
	Good	6.426*
	Poor	2.872

Note: LPC = least preferred co-worker; LMR = leader member relations.

\*  $p < .05$ .

\*\*  $p < .01$ .

the hypothesized direction. Thus, hypothesis 1b was supported by the data.

To test hypothesis 2 each of the judges' scores on the unstructured task on each criterion and on the total score was converted to a  $T$  score. The median  $T$  score for the three judges on each of the criteria and for the total score was the dependent variable. To analyze these data a  $2 \times 2$  factorial analysis of variance was run between LPC and LMR with 14 cases per cell. The means of these data are presented in Table 2, and the results of the analysis of variance are presented in Table 3. An interaction effect was hypothesized with high LPC leaders being more effective when the relations are good, and low LPC leaders being more effective when they are poor. Table 2 shows that the means of the total score and all the criteria except adequacy of recommendation were in the predicted direction. For adequacy of recommendation only the groups with good LMR were in the predicted direction. Table 3 shows that for the total score and when the criterion was style, the interaction effect was significant. When the criterion was persuasibility, the interaction effect approached significance. Thus, there is moderate support for hypothesis 2.

To test hypothesis 2a, which investigated Cell 4 of the contingency model, individual comparisons on the unstructured task when the LMR were good were made between the high and low LPC leaders on each of the criteria and on the total score. It was hypothesized that high LPC leaders are more effective than low LPC leaders. The results of these analyses are presented in Table 4, and the means are presented in Table 2. For the total score and when the criteria were adequacy of recommendation and persuasibility, the results were significant. When the criterion was style, the results approached significance. All means were in the hypothesized direction. Since style was in the predicted direction and since the rest of the analyses were significant, it was concluded that hypothesis 2a was supported by the data, and that Cell 4 was replicated.

To test hypothesis 2b, which investigated Cell 8 of the contingency model, individual comparisons on the unstructured task when the LMR were poor were made between the high and low LPC leaders on each of the criteria and on the total score. The results of these analyses are presented in Table 4, and the means are presented in Table 2. When the criterion was style, the  $F$  test ( $df = 1, 52$ ) was significant. The means were all in the hypothesized direction, except adequacy of recommendation. Thus, hypothesis 2b and Cell 8 of the model received only limited support.

#### D. DISCUSSION

The results of this study offer partial support for the contingency model. Cell 8 received limited support, Cells 4 and 6 were strongly supported, and Cell 2 was not supported by the data. Further, the data concerning Cell 2 were in the reverse direction, which is consistent with the results of the study by Hardy (11) and several other researchers. This also adds credence to the already discussed argument that in educational settings power is a more important determinant of situation favorability than is task structure. If power is a more important determinant of situation favorability than is task structure, and Cells 2 and 3 reverse position, the same should occur between Cells 6 and 7, with Cell 6 being a more favorable cell for low LPC leaders and Cell 7 for high LPC leaders. However, this was not consistent with the data found in this study concerning Cell 6, in which high LPC leaders were more effective than low.

Another explanation for the failure of this study to replicate Cell 2 concerns LMR. If those group members in Cell 2 did not actually have good LMR, the results should be consistent with those obtained in Cell 6, which they approached. This explanation, however, was rejected for three reasons.

First, the results of both the group atmosphere scale and the group evaluation scale indicated that the groups did have good LMR. Second, the sociometric data collected prior to the experimental session indicated that the groups should have had good relations. Third, if the LMR were poor, then when the groups did the unstructured task, the means should have been in the same direction as were the data collected to test Cell 8. Since they were not in this direction, and since the premeasure and postmeasures indicated that the LMR were good, this explanation to account for the failure to replicate Cell 2 of the model was rejected.

Another possible explanation for this study's failure to replicate Cell 2 would be the improper manipulation of position power. If the groups actually possessed strong power instead of weak power, the cell being tested would have been 1 instead of 2. Further, if both task structure and position power were improperly manipulated, the cell being tested would have been 3 instead of 2. In both of these instances low LPC leaders would have been favored over high LPC leaders, which was contrary to the data. Further, the results of the position power checklist indicated that the leaders did not have strong position power. Thus, improper manipulation of position power was rejected as an explanation for the failure of this study to replicate Cell 2 of the model.

If task structure was improperly manipulated, as discussed by Hardy (11), and the structured task was actually unstructured, the cell being tested would have been Cell 4, where high LPC leaders would have been favored, instead of Cell 2. This was consistent with the data. However, if task structure was improperly manipulated, not only would Cell 2 have been mistested as Cell 4, but also Cell 6 with poor relations and a structured task would have been mistested as Cell 8. If this were the case, low LPC leaders would have been favored over high. This was not consistent with the data, since high LPC leaders were more effective than low when Cell 6 was tested. Thus, task structure was assumed to have been properly manipulated, and this explanation to account for the failure of Cell 2 to be replicated was rejected.

This study is just one more which adds confusion to the issue surrounding the difficulty in replicating Cell 2 of the contingency model. Two of the possible explanations postulated by Hardy (11) and discussed earlier have been rejected. Some alternative explanation is needed. It might be well to direct future investigations to the composition of group members in relation to situation favorability. Further, it seems that more meaningful research would be obtained if the situation dimensions were looked at in some other fashion than dichotomizing them.

It should also be noted that in this study no main effects on the LMR



dimension existed. Groups did not significantly differ in respect to good or poor relations on either task. This consistency across this favorability dimension is in agreement with the findings of Csoka and Fiedler (4) and Chemers and Skrzypek (3), but it is inconsistent with the findings of Shiflett and Nealey (20).

The results of this study offered partial support for the contingency model when it was applied to intact junior high school classrooms. In particular, Cells 4 and 6 were supported, and Cell 2 was not supported by the data. In addition Cell 8 received only minimal support. Further research, however, is needed if the contingency model is to become a useful tool to aid classroom teachers in maximizing group output in their classes. To better account for the discrepancies that exist between various studies investigating the contingency model, and in particular to better account for the failure of many studies to replicate Cell 2, future experimentation needs to be concerned with developing more precise data and methodology indicating the degree of favorability along with the use of more precision in experimental analysis.

#### REFERENCES

1. ANDERSON, L. R. Some effects of leadership training on intercultural discussion groups. Unpublished Doctoral dissertation, University of Illinois, Urbana, 1964.
2. BUTTERFIELD, D. A. Supervisory behavior and effectiveness: A review of empirical research and current thinking. Mimeographed paper, Department of Psychology, University of Michigan, Ann Arbor, 1967.
3. CHEMERS, M. M., & SKRZYPEK, G. J. An experimental test of the contingency model of leadership effectiveness. *J. Personal. & Soc. Psychol.*, 1972, **24**, 172-177.
4. CSOKA, L. S., & FIEDLER, F. E. The effect of leadership experience and training in structured military tasks: A test of the contingency model. *Organiz. Behav. & Hum. Perform.*, in press.
5. FIEDLER, F. E. A contingency model of leadership effectiveness. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 1). New York: Academic Press, 1964. Pp. 150-190.
6. ———. The effect of leadership and cultural heterogeneity on group performance: A test of the contingency model. *J. Exper. & Soc. Psychol.*, 1966, **2**, 237-264.
7. ———. *A Theory of Leadership Effectiveness*. New York: McGraw-Hill, 1967.
8. ———. Personality, motivational systems, and behavior of high and low LPC persons. Technical Report No. 70-12, Organizational Research Laboratory, University of Washington, Seattle, 1970.
9. ———. Predicting the effects of leadership training and experience from the contingency model. *J. Appl. Psychol.*, 1972, **56**, 114-119.
10. GRAEN, G., ALVARES, K. M., ORRIS, J. B., & MARTELLA, J. A. Contingency model of leadership effectiveness: Antecedent and evidential results. *Psychol. Bull.*, 1970, **74**, 285-296.
11. HARDY, R. C. Effects of leadership style on the performance of small classroom groups: A test of the contingency model. *J. Personal & Soc. Psychol.*, 1971, **19**, 367-374.
12. HILL, W. The validation and extension of Fiedler's theory of leadership effectiveness. *Acad. Manag. J.*, 1969, March, 33-47.

13. HUNT, J. G. A test of Fiedler's leadership contingency model in three organizational settings. Doctoral dissertation, University of Illinois, Urbana, 1966.
14. MITCHELL, T. R., BEGLEN, A., ONCHEN, G., & FIEDLER, F. E. The contingency model: Criticism and suggestions. *Acad. Manag. J.*, 1970, 13, 253-267.
15. O'BRIEN, G. E., FIEDLER, F. E., & HEWETT, T. The effects of programmed cultural training upon the performance of volunteer medical teams in Central America. Office of Naval Research Technical Report No. 76, Group Effectiveness Research Laboratory, University of Illinois, Urbana, 1969.
16. REILLY, A. J., III. The effects of leadership styles on group performance: A field experiment. Published Doctoral dissertation, Industrial Relations Center, Ames, Iowa, 1968.
17. SHAW, M. E. Scaling group tasks: A method for dimensional analysis. Clearinghouse, Springfield, Virginia, 1963.
18. SHAW, M. E., & BLUM, J. M. A comparison of two types of leadership in communication networks. *J. Abn. & Soc. Psychol.*, 1966, 50, 127-134.
19. SHIPLETT, S. C. The contingency model: Some implications of its statistical and methodological properties. United States Medical Research Laboratory, Fort Knox, Kentucky, 1972.
20. SHIPLETT, S. C., & NEALEY, S. M. The effects of changing leader power: A test of "situation engineering." *Organiz. Behav. & Hum. Perform.*, 1972, 7, 317-382.
21. TVERSKY, A., & KAHNEMAN, D. Belief in the law of small numbers. *Psychol. Bull.*, 1971, 76, 105-110.
22. WHAT DO I KNOW ABOUT BENNY? Film in the series *Critical Moments in Teaching*. New York: Holt, Rinehart & Winston, 1967.
23. WINER, B. J. *Statistical Principles in Experimental Design*. New York: McGraw-Hill, 1962.

*Institute for Child Study*  
*University of Maryland*  
*College Park, Maryland 20742*

## IS TEST ANXIETY RISING?

*Lehman College and Finch College*

---

ARLINE L. BRONZAFT AND GILDA F. EPSTEIN<sup>1</sup>

---

### SUMMARY

Test anxiety scores of students at three New York metropolitan colleges were compared with test anxiety scores obtained at these colleges five years earlier. At all three schools mean facilitating test anxiety scores had decreased, whereas mean debilitating test anxiety scores had increased. It was suggested that changes in the student populations during the years might account for the changes in test-taking attitudes.

---

### A. INTRODUCTION

The literature on the relationship between test anxiety and college performance is replete with contradictions and disagreements. Some investigators (1, 4, 7) have reported significant relationships between test anxiety and college achievement, whereas others (5, 9) have found no relationship. Pervin (7) suggested that such contradictory findings may result from administering tests at different times and under different conditions. Bronzaft and Epstein (3) failed to replicate an earlier finding of a significant relationship between test anxiety and college performance in socially mobile male students, and suggested that the higher level of debilitating anxiety in the later sample might have accounted for the discrepant findings. The Bronzaft and Epstein study raised the question of a possible change in test-taking attitudes and feelings over the years. Yet, a recent study by Peterson (8) reported no change in test-taking attitudes for three midwestern colleges tested at two different periods, 1962-63 and 1968-69.

The present study reports test anxiety levels of students at three colleges in which Bronzaft (2) had previously measured test anxiety in 1965.

### B. METHOD

The Alpert-Haber (1) Achievement Anxiety Test (AAT), which consists

---

\* Received in the Editorial Office on May 14, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was supported, in part, by a grant from the Finch College Faculty Fund to the second author.

of a facilitating test anxiety scale (AAT+) and a debilitating test anxiety scale (AAT—), was administered to 69 male students at College A, 150 at College B, and 98 at College C. All three colleges are located in the New York metropolitan area. Students were tested during 1970-72.

### C. RESULTS

Mean facilitating test anxiety (AAT+) scores and mean debilitating test anxiety (AAT—) scores were obtained for the 1970-72 samples and compared to the 1965 samples. To determine whether there were differences between the two sets of means at each college, *t* tests were done. Means, standard deviations, and *ts* are presented in Table 1.

TABLE 1  
MEANS AND STANDARD DEVIATIONS OF SCORES ON THE ALBERT-HABER ACHIEVEMENT ANXIETY TEST AND DIFFERENCES BETWEEN MEANS

Sample	N	1965		N	1970-72		t
		M	SD		M	SD	
Facilitating test anxiety scores (AAT+)							
College A	317	26.49	4.36	69	23.83	4.60	4.43
College B	314	25.86	4.67	150	23.68	4.91	4.54
College C	209	25.94	4.32	98	23.81	4.65	3.80
Debilitating test anxiety scores (AAT—)							
College A	317	26.75	5.46	69	30.65	5.79	5.06
College B	314	27.60	6.16	150	28.97	6.56	2.54*
College C	209	26.77	6.28	98	29.00	5.61	3.09

\* This value significant at  $p < .05$ , all others  $p < .01$ .

The mean AAT+ scores for the 1970-72 samples at all three colleges were significantly smaller than those for the 1965 samples. The mean AAT— scores for the 1970-72 samples were significantly larger than those obtained in 1965.

### D. DISCUSSION

Male college students tested at three different colleges in 1970-72 had more negative and less positive test-taking attitudes than students tested at the same colleges in 1965. The nature of the student population in 1965 and 1970 might explain the change in attitude. By 1970 the students tested probably represented a wider range of socioeconomic and academic backgrounds than those tested in 1965. This was most clearly true of the sample drawn from the City University of New York where an open admissions policy went into effect in 1970.

Since the three colleges in the present study were all located in the New York metropolitan area, it is possible that the increase in debilitating test

anxiety and decrease in facilitating test anxiety may be a local phenomenon. In fact, ATT scores obtained from 220 University of Virginia male students in 1971, as reported in an unpublished study (6), showed a mean AAT— score of 25.30 and a mean AAT+ score of 25.71. These scores are more in line with scores reported during the 1960s than with those reported in the present study. However, since earlier test anxiety scores at the University of Virginia were not available, it remains possible that these scores may represent an increase in debilitating test anxiety or a decrease in facilitating test anxiety. Furthermore, the problem of interpreting the findings of Peterson (8) which indicated no change in test anxiety levels is compounded by the fact that test anxiety in that study was assessed by means of a projective test rather than by a questionnaire instrument, such as the AAT.

## REFERENCES

1. ALPERT, R., & HABER, R. N. Anxiety in academic achievement situations. *J. Abn. Soc. Psychol.*, 1960, **61**, 207-215.
2. BRONZAFI, A. L. Test anxiety, social mobility, and academic achievement. *J. Soc. Psychol.*, 1968, **75**, 217-222.
3. BRONZAFI, A. L., & EPSTEIN, G. F. Test anxiety, academic achievement, and the open admissions student. *J. Soc. Psychol.*, in press.
4. DEMBER, W. M., NAIRNE, F., & MILLER, F. J. Further validation of the Alpert-Haber achievement anxiety test. *J. Abn. & Soc. Psychol.*, 1962, **65**, 427-428.
5. GROOMS, R. R., & ENDLER, N. S. The effect of anxiety on academic achievement. *J. Educ. Psychol.*, 1960, **51**, 299-304.
6. HARRISON, P. H. Socioeconomic factors, test anxiety and academic achievement. Unpublished study, Finch College, New York, New York, 1971.
7. PERVIN, L. A. Aptitude, anxiety and academic performance: A moderator variable analysis. *Psychol. Rep.*, 1967, **20**, 215-221.
8. PETERSON, V. O. A clinical and longitudinal study of anxiety and attitudinal trends in college students using a projective test as a disguised technique. *Diss. Abst. Internat.*, 1972, **32**, (8-B), 4868-4869.
9. SARASON, S. B., & MANDLER, G. Some correlates of test anxiety. *J. Abn. & Soc. Psychol.*, 1952, **47**, 810-817.

Department of Psychology  
Herbert H. Lehman College of the  
City University of New York  
Bronx, New York 10468





## THE ASTROLOGICAL "THEORY" OF PERSONALITY: AN UNBIASED TEST BY A BIASED OBSERVER\*<sup>1</sup>

*California State University, San Jose*

---

ROBERT J. PELLEGRINI

---

### SUMMARY

The astrological assumption of personality differences among individuals born under different sun signs was investigated. Significant differences were obtained between sun sign groups on four of the 18 scales of the California Psychological Inventory (CPI) as follows: Communality, Socialization, Flexibility, and Femininity. The overall pattern of results based on *F* ratios, multiple comparisons between means, and estimates of strength of association was interpreted as showing no meaningful relationships between sun sign and scores on Communality, Socialization, and Flexibility. Adopting even the most conservative standards of inference, however, all of the test statistics employed indicated a powerful effect on the Femininity scale. Implications for future research were discussed.

---

### A. INTRODUCTION

Ironically, there has been practically no research by psychologists into one of the oldest and most popular "theories" of personality: namely, astrology. Perhaps we have been reluctant to risk dignifying as a legitimate area of investigation what most psychologists, the writer included, have long considered mystical nonsense. But ignored or not, astrology has never gone away. To the contrary, it seems to be flourishing among as zealous a group of followers as ever after having been rediscovered in the recent resurgence of interest in the occult. Most unfortunately, the lack of formal inquiry has left unchallenged the conjecture and phantasy so often presented as reliably established fact by the purveyors of astrological wisdom. The study reported here was designed to remove a central assumption of astrology from the protective cover of unavailable data.

---

\* Received in the Editorial Office on May 14, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was supported in part by a grant from the Center for Research and Advanced Studies, California State University, San Jose. The author wishes to thank Mr. Bruce Antal for his assistance in the collection of data for this study.

Essentially, astrology is built around the idea that personality and destiny are unalterably, and hence predictably, predetermined by the position of the sun, moon, and planets at the moment of birth (e.g., 5, 12). Each of the so-called "signs of the zodiac" (Taurus, Leo, etc.) is associated with a set of characteristics which are, presumably, transmitted to those people born under its influence. Prediction thus resolves to a calculation of the position of the sun, moon, and planets at the time of one's birth, and the explication of their significance for his life through the artistry of astrological interpretation.

An individual's "sign" or "sun sign" is determined by the position of the sun in the zodiac—a giant circle which encompasses the yearly path of movement of the sun, moon, and planets—at the time of his birth. Since our calendar is based on the laws of solar and lunar motion, the signs can be ascertained easily from birthdates. So, for example, all people born while the sun is crossing the zodiac zone called Aries (March 21 to April 20) are Aries; those born from April 21 to May 21 are born under the sign of Taurus; and so on.

Since the relative positions of solar bodies are changing constantly, the most accurate predictions, supposedly, require a complete horoscope—a thorough charting of the skies at the precise moment of a child's birth. But the sun sign, even though it constitutes only a part of the horoscope, is generally acknowledged by astrologers as the critical component [see Goodman (5), p. xviii].

The wondrous insights promised to the prospective reader of sun sign astrology are exemplified in the following cover lines from a recent best-seller (5):

Is your *husband* your kind of man? How do you get him and keep him?  
Is your *wife* passionate, capable, suspicious, extravagant, submissive? What  
are your wife's secrets? What does she *really* want and need? What makes  
your *lover* want you? How do you make him want you more? Where is  
your *child* going and how can you help him get there? Why did your *boss*  
hire you, and how can you keep your job? How do you make your *employee*  
come in early, stay late, and love it?

Clearly, many of the writings by astrologers are offered as sources for instant revelations about human nature. But astrological premises have not been confirmed by the few empirical studies that have been conducted in this area. Farnsworth's (3) data, for example, failed to support the contention that people born under the sign of Libra are more likely to be characterized by outstanding artistic talent. Similarly, the work of Bok and Mayall (1)

revealed no significant relationship between astrological sign and the occupation of particular professional roles. Consistent with the results of earlier studies, Silverman (11) found no differences among natives of the 12 signs in their expressed values, nor did his data confirm astrological predictions about marital compatibility of members of the various signs.

Although there is variability between writers in the specific characterizations given, *all* astrological analysis assumes personality differences between people born under different signs. With the exception of Silverman's (11) research, however, there have been no attempts to test for relationships between sun sign and individual differences on objective measures of personality. In view of the apparent popularity of astrology, more detailed inquiry into the validity of its central assumption by such methods seems warranted. Along these lines, the present study tested the general hypothesis of differences between sun sign groups on the dimensions of personality measured (6) by the California Psychological Inventory (CPI). The CPI was selected as the test instrument on the basis of two primary considerations: (a) the desirability of testing for possible differences over a fairly wide range of characteristics, and (b) the social emphasis of astrological descriptions of personality. The CPI provides a profile of scores from scales constructed to measure a variety of interpersonal response traits, and was thus particularly appropriate for the purposes of the study.

## B. METHOD

### 1. *Procedure*

In order to ensure an adequate representation of individuals from each of the 12 sun signs, the CPI was administered to large groups of undergraduates at the California State University, San Jose. Sun signs were determined from birthdates listed on records filed in the University's Registrar's office. In classifying birthdates by sun sign, a problem arises with regard to the so-called "cusp" dates—those falling within the first or last four days of a zodiacal month. Although some astrologers (e.g., 5) minimize the importance of the cusp, others insist that people born on cusps are influenced by two signs rather than one, and should be considered separately. So as to approximate more closely the standards of the astrological purist, the CPI profiles from individuals whose birthdates fell on cusps were eliminated.

### 2. *Test Sample*

After incomplete and cusp CPI records were discarded, the test sample was selected from the remaining 412 usable profiles so as to include the

largest equal number of male and female members of each sun sign grouping. The final sample was thus comprised of 288 students, 144 males and 144 females, with 12 males and 12 females chosen at random from each of the 12 sun sign groupings in the larger pool.

### C. RESULTS

The astrological assumption of personality differences between sun sign groupings was tested by submitting the data from each of the 18 CPI scales to a 2 (Sex)  $\times$  12 (Sun Sign) completely randomized factorial analysis of variance with 12 scores per cell. Neither the main effect on the sex factor nor the interaction was significant in any of these analyses. The main effect on the sun sign variable, however, was significant in four of the 18 analyses.

In order to evaluate rigorously the effects of principal interest, strength of association (7) between sun sign grouping and scores on all of the CPI scales was measured by calculating estimates of omega squared ( $\omega^2$ ). The  $F$  ratios, corresponding probabilities, and  $\omega^2$  values for the main effect on the sun sign factor are summarized in Table 1. As shown, significant  $F$

TABLE 1  
SUMMARY OF ANALYSES OF DIFFERENCES BETWEEN SUN SIGN GROUPS ON  
ALL 18 CALIFORNIA PSYCHOLOGICAL INVENTORY (CPI) SCALES

Scale	$F$	est. $\omega^2$
Dominance	1.16	.01
Sense of Well-Being	<1.00	.00
Tolerance	<1.00	.00
Intellectual Efficiency	<1.00	.00
Good Impression	<1.00	.00
Capacity for Status	1.43	.02
Responsibility	1.77	.03
Communality	4.50**	.12
Psychological Mindedness	<1.00	.00
Sociability	1.76	.03
Socialization	2.29*	.05
Achievement via Conformance	1.38	.02
Flexibility	2.24*	.05
Social Presence	1.73	.03
Achievement via Independence	1.10	.00
Femininity	24.24**	.47
Self-Acceptance	1.84	.03
Self-Control	<1.00	.00
Median		.015

Note: All  $p$  values based on  $df = 11/264$ .

\*  $p < .025$ .

\*\*  $p < .001$ .



ratios were obtained in the analyses of the following scales: Communality, Socialization, Flexibility, and Femininity. Duncan's range test was used to conduct multiple comparisons between pairs of means for each of these significant effects. The  $F$  ratios for the Socialization and Flexibility scales were only marginally ( $p < .05$ ) significant, so it is not surprising that Duncan's test, a more conservative procedure than the omnibus  $F$  test, revealed just one barely significant comparison on the former, and none on the latter. But even with the more substantially significant ( $p < .025$ )  $F$  obtained on Communality, only three of the 66 pairwise comparisons were significant.

The  $\omega^2$  values provide the most appropriate index for overall appraisal of these results for two reasons. First, the actual probability of Type I error may have exceeded the estimate of the stated level ( $p < .05$ ) because of the large number of  $F$  tests conducted. Second, the design employed allowed for such a powerful test of the main effect on the sun sign factor that even relatively trivial differences might have produced a significant  $F$ . The  $\omega^2$  statistic indicates not significance, but the proportion of total uncertainty about the dependent variable reduced by knowledge of the independent variable (7). The  $\omega^2$  is thus one technique for evaluating the meaningfulness of a statistical difference.

Given the foregoing rationale, the crucial summary statistic in the study is the average  $\omega^2$  for all scales of the CPI combined. As shown in Table 1, the median (the most representative measure of central tendency for this markedly skewed distribution)  $\omega^2$  value was less than .02, indicating essentially no relationship between astrological zodiac sign and the CPI scales *taken as a group*. The significant  $F$  ratios on the Communality, Socialization, and Flexibility scales are, therefore, most parsimoniously interpreted as instances of Type I error in view of their respectively inconsequential  $\omega^2$  values of .12, .05, and .05. This interpretation is reinforced by the almost complete absence of significant pairwise contrasts between means on these variables.

The dramatically significant  $F$  and the 36 highly significant differences between pairs of means on the Femininity scale, however, cannot be attributed to chance quite so easily. As shown in Table 2, half of the sun sign groupings—Capricorn (Cap), Sagittarius (Sa), Libra (Li), Leo (Le), Scorpio (Sc), and Virgo (Vi)—scored distinctly higher on the Femininity scale than did the other six groupings—Gemini (Ge), Taurus (Ta), Pisces (Pi), Aquarius (Aq), Aries (Ar), and Cancer (Can). The clarity of separation between groupings on the Femininity variable is demonstrated by the fact

TABLE 2  
SUMMARY OF MEAN DIFFERENCES BETWEEN SUN SIGN GROUPS<sup>a</sup> ON FEMININITY SCALE  
OF CALIFORNIA PSYCHOLOGICAL INVENTORY

Sign	Mean	Can	Ar	Aq	Pi	Ta	Ge	Vi	Sc	Le	Li	Sa	Cap
Can	15.96		.33	.41	.96	1.46	1.62	6.41*	7.00*	7.00*	7.50*	7.91*	8.50*
Ar	16.29			.08	.63	1.13	1.29	6.08*	6.67*	6.67*	7.17*	7.58*	8.17*
Aq	16.37				.55	1.05	1.21	6.00*	6.59*	6.59*	7.09*	7.50*	8.09*
Pi	16.92					.50	.66	5.45*	6.04*	6.04*	6.54*	6.95*	7.54*
Ta	17.42						.16	4.95*	5.54*	5.54*	6.04*	6.45*	7.04*
Ge	17.58							4.79*	5.38*	5.38*	5.88*	6.29*	6.88*
Vi	22.37								.59	.59	1.09	1.50	2.09
Sc	22.96									.00	.50	.91	1.50
Le	22.96										.50	.91	1.50
Li	23.46											.41	1.00
Sa	23.87												.59
Cap	24.46												

Note: High scores indicate more feminine interests.

<sup>a</sup> Can = Cancer, Ar = Aries, Aq = Aquarius, Pi = Pisces, Ta = Taurus, Ge = Gemini, Vi = Virgo, Sc = Scorpio, Le = Leo, Li = Libra, Sa = Sagittarius, Cap = Capricorn.

\* Mean differences established as significant ( $p < .001$ ) by Duncan range test.

that all possible pairwise comparisons between the two sets of groupings were highly significant ( $p < .001$ ), while none of the Duncan contrasts within sets were significant. (See Table 1, where the  $\omega^2$  of .47, reflecting an estimate of 47% of the variance in Femininity scores accounted for by knowledge of zodiac sign, supports these indications of a noteworthy effect.)

#### D. DISCUSSION

For the most part, the data obtained here were consistent with the results of previous research which have failed to confirm the fundamental astrological hypothesis of personality differences among natives of the 12 signs of the zodiac. In general, there were no meaningful relationships between sun signs and the personality traits measured by the CPI. The important exception, however, was a remarkably strong association between sun sign and scores on the CPI scale measure of Femininity.

High scorers on the Femininity scale are described in the CPI manual (6) as appreciative, patient, helpful, gentle, moderate, persevering, sincere, accepting of others, conscientious, and sympathetic. Low scorers are described as outgoing, hard-headed, ambitious, masculine, active, robust, restless, manipulative, opportunistic, and impatient. It would be of interest to know how closely such adjectives, applied to sun sign categories according to their score distribution on the Femininity scale, correspond to astrological characterizations of those signs. But perhaps this is a matter best left to astrologers.

The powerful effect on the Femininity variable is very puzzling. It would require an almost absurdly conservative approach to data interpretation to attribute to chance a relationship of such magnitude. And yet, to the writer's knowledge, there is just no way to explain it in terms of established principles of any orthodox theory of personality. Contemporary students of personality might dispute the appropriateness of defining as "femininity" the attributes associated with high scores on the CPI scale so-designated by Gough (6). But further research is clearly indicated into possible differences among people born at different times of the year in the specific traits that contribute to scores on the scale as it is currently used.

On the assumption that the effect on the femininity variable is a reliable one, the problem of how to account for it is challenging indeed. Why should people born at different times of the year differ in *any* dimension of personality? It is remotely possible that differences of this sort, should they exist, may be attributable to quite "unearthly" causes. A plausible, and decidedly nonoccult explanation might be conjectured from the standpoint of a developing astrology based on traditional methods of inquiry. This "modern astrology" seeks to identify reliable, objective relationships between astronomical phenomena and biochemical processes. One of the most provocative hypotheses identified with this approach is a theory relating moon phases and planetary positions to human fertility, as well as the sex and health of offspring (9). Somewhat less controversial is the literature on biological effects of electromagnetic activity associated with solar and lunar events (e.g., 2). There has even been some research into the psychiatric importance of such geomagnetic activity (4, 10). From this point of view, it might be speculated that certain cyclical, geomagnetic events may somehow so affect reproduction that people born at different times of the year are constitutionally predisposed to develop particular traits of personality.

At least one psychologist (8) has suggested that a scientifically based astrology can provide a more sophisticated and effective approach to understanding personality than any of the existing systems. This remains to be seen. But few psychologists would argue that all of the sources contributing to intra- and interindividual differences in personality have been clearly identified. To deny *a priori* the value of novel lines of inquiry is to be guilty of a naive obscurantism or afflicted by an unenviably narrow perspective. It is not inconceivable that human biological and psychological processes may be affected significantly by cosmic-atmospheric conditions associated with astronomical events. The powerful role of the environment in the develop-

ment of personality is well-established. Perhaps the new astrology will provide the impetus for serious consideration of geomagnetic influences on personality—environmental parameters which have been almost totally unexplored. And it is not entirely unlikely that increased interest in relationships between psychological and cosmobiological variables could stimulate a genuinely productive line of research. But that too remains to be seen.

#### REFERENCES

1. BOK, B. J., & MAYALL, M. W. Scientists look at astrology. *Sci. Monthly*, 1941, 52, 223-244.
2. BROWN, F. A., JR. A hypothesis for extrinsic timing of circadian rhythms. *Can. J. Botany*, 1969, 47, 287-298.
3. FARNSWORTH, P. Aesthetic behavior and astrology. *Charac. & Personal.*, 1937, 6, 335-340.
4. FRIEDMAN, H., BECKER, R. O., & BACHMAN, C. H. Psychiatric ward behaviour and geophysical parameters. *Nature*, 1965, 205, 1050-1052.
5. GOODMAN, L. Sun Signs. New York: Bantam, 1968.
6. GOUGH, H. G. California Psychological Inventory. Palo Alto, Calif.: Consult. Psychol. Press, 1956-1957.
7. HAYS, W. L. Statistics for Psychologists. New York: Holt, Rinehart & Winston, 1963.
8. METZNER, R. Astrology: Potential science and intuitive art. *Aquarian Agent*, 1970, 1, 3-6.
9. OSTRANDER, S., & SCHROEDER, L. Astrological Birth Control. Englewood Cliffs, N. J.: Prentice-Hall, 1972.
10. POKORNY, A. D., & MEFFERD, R. B., JR. Geomagnetic fluctuations and disturbed behavior. *J. Nerv. & Ment. Disor.*, 1966, 143, 140-151.
11. SILVERMAN, B. Studies of astrology. *J. of Psychol.*, 1971, 77, 141-149.
12. STEARN, J. A Time for Astrology. New York: Signet, 1972.

*Department of Psychology*

*California State University, San Jose*

*San Jose, California 95114*



## FAILURE TO FIND RELATIONSHIPS BETWEEN FAMILY CONSTELLATION AND ACHIEVEMENT MOTIVATION\*

*University of Port Elizabeth, South Africa*

---

D. J. W. STRÜMPFER

---

### SUMMARY

Ss were 158 male and 160 female university students. The Holmes-Tyler self-peer rating and Mehrabian resultant measures of achievement motivation showed no significant relationships to ordinal position, family size and density, or sex/ordinal position dyads. Family size, as well as a family size and density index, showed significant interaction with father's occupational status in three out of eight analyses of variance involving occupational status. Some characteristics of the family size and density index were discussed.

---

### A. INTRODUCTION

Studies on the relationship between family constellation and achievement motivation typically use birth order and family size as variables. Atkinson and Miller (*cf.* 5) and Sampson (15) found firstborns to have higher achievement motivation than later borns, but other authors (9, 12, 14, 23) found no such relationship. Rosen (12) found that the influence of family size on achievement motivation varied with social class, while Rosenfeld (14) found complex interactions between sex, ordinal position, and stimulus items. Oberlander, Jenkin, Houlihan, and Jackson (11) studied achievement test scores and scholastic achievement: firstborns obtained higher *IQ* scores than later borns; family size was not significantly related to any of their measures, but a birth order  $\times$  family size interaction was found to be more meaningfully related to them than either birth order or family size alone.

Other aspects of family constellation, however, could also be investigated in this context. One such aspect would be spacing of children in the family. Waldrop and Bell (20) developed a family size and density index based on weighted scores for number of children, time span to next younger sibling,

---

\* Received in the Editorial Office on May 15, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.



time span to next older sibling, and average time span between births. These authors (19, 20) used the index only in studies of nursery school children and newborn infants.

An aspect of family constellation which has been neglected in connection with achievement motivation is that of sibling interaction. Sutton-Smith *et al.* (17, 18) classified Ss from two-child families into sex/ordinal position dyads. With M or F denoting the sex of S or sibling and 1 or 2 denoting first- or second-born, the following eight dyads were described: M1M, MM2, M1F, FM2, F1F, FF2, F1M, and MF2. Other authors (2, 8) have used similar systems, with age-spacing limits added in the selection of Ss. Analysis of 24 triads of three-child families (13) requires exceptionally large samples.

In the present study the usual comparisons between firstborns and later borns were made, but also comparisons between only, first, youngest, and middle children. Waldrop and Bell's (20) index and its four components separately, as well as comparisons of sibling dyads, were also used. The achievement motivation measures were those of Holmes and Tyler (7) and Mehrabian (10).

## B. METHOD

### 1. Subjects

Ss were 158 male and 160 female students at the University of Port Elizabeth and Rhodes University. The majority of them were in Introductory Psychology, thus including a fairly representative sample of first-year Arts and Social Science students; a sample of first-year students in Economic Sciences was also included. All Ss were white. Fathers' occupations were classified according to Hall and Jones' (4) seven-class social grading of occupations. Fathers' occupations were reported by 107 men and 106 girls; the mean occupational levels were, respectively, 2.27 ( $SD = 1.29$ ) and 2.16 ( $SD = 1.24$ ). However, the distributions were skewed, with higher occupations overrepresented (in Classes 4, 5, and 6 only 11 in male and 16 in female sample). The mean educational achievement of fathers was 12.55 years ( $SD = 2.78$ ) for 125 men and 12.92 years ( $SD = 2.83$ ) for 138 girls; for the mothers it was 11.65 years ( $SD = 1.95$ ) for 130 men and 12.08 years ( $SD = 1.95$ ) for 125 girls.

The data were gathered during regular class periods, with a request to participate in research, and reassurance of confidential treatment of information.

All calculations were carried out separately for males and females, since

achievement motivation may have different sex/ordinal position relationships in the two sexes (15).

## 2. Achievement Motivation Measures

Holmes and Tyler (7) developed a self-peer rating procedure in which *S* lists 10 of his friends; he then rates himself in relation to each friend, first, in terms of the need to achieve as defined in the instructions, and then in terms of working longer and harder on an academic task. This measure was selected for its simplicity, as well as its validity in comparison with that of projective measures (6, 7). In the present study split-half reliabilities [Guttman formula, see Anastasi (1)] for the two separate ratings were too low to warrant further analyses for them separately. The sum of the two scores for each *S* showed reliabilities of, respectively, .60 and .75 in the male and female groups; this total score (HT) was used.

Mehrabian (10) constructed an inventory of resultant achievement motivation (RAM) to distinguish high achievers with a stronger motive to achieve than to avoid failure, from low achievers with a stronger motive to avoid failure than to achieve. Items were written in such a way that response to each indicates a behavioral disposition that has been found to be characteristic of high or low achievers. The advantage of the RAM is that it provides a single, resultant measure of achievement motivation, rather than the common approach of using a projective measure in combination with a test anxiety questionnaire. Separate forms are used for males and females. In the present study the RAM showed split-half reliabilities of .69 and .55 for the male and female groups.

In the subsample of male Economic Sciences students end-of-the-year marks (%) were used to determine the predictive validity of the HT and RAM, administered six months before. The HT showed *rs* of .28 with Economics ( $n = 55$ ) and .39 with Mercantile Law ( $n = 32$ ); these *rs* were significant beyond the .05 level, while those with Business Economics and Accountancy were below this level. Among these four subjects only Accountancy correlated significantly with the RAM; viz., .34 ( $n = 34$ ,  $p < .05$ ).

Significant *rs* with the Crowne-Marlowe Social Desirability Scale (3) were found for the RAM in a female subsample ( $r = .28$ ,  $n = 115$ ,  $p < .01$ ) and for both the HT and RAM in a male subsample ( $r = .23$  and  $.21$ ,  $n = 119$ ,  $p < .05$ ).

In the total male and female samples the *rs* between HT and RAM were, respectively, .28 ( $p < .001$ ) and .23 ( $p < .01$ ).

## C. RESULTS

### 1. Ordinal Position

In the male sample the distribution of the Ss between the various ordinal positions was as follows: single = 10, first = 62, youngest = 39, and middle = 47. In the female sample the distribution was as follows: single = 14, first = 50, youngest = 50, and middle = 46. Single and firstborn Ss were kept separate on the basis of Warren's (21) indications that they should be differentiated. Both groups showed the relative preponderance of firstborn in university samples frequently reported before (21).

One-way analyses of variance comparing only children, firstborn, youngest, and middle children with respect to HT and RAM scores did not yield significant *F* values in either the male or female samples. Comparing firstborns plus only children with later borns also yielded nonsignificant *t* values.

### 2. Family Size and Density

The four components of the Waldrop-Bell index were each correlated with HT and RAM for the male and female samples separately. None of these *rs* even approached significance, the highest of them being .12.

The interrelations of these four family variables have to be considered carefully. Waldrop and Bell (20) intercorrelated them and finding high *rs*, they extracted one factor, using Hotelling's principal components method. They pointed out that the *rs* between the average time between births and the time spans to next younger and next older siblings were part-whole *rs*, so that the contributions of the latter two variables to the factor were somewhat overestimated. In their study these part-whole *rs* were .43 and .79, while in the present study they were .51 and .82 in the male sample, and .70 and .80 in the female sample. A large discrepancy between their findings and those of the present study, however, occurred with respect to the four other intercorrelations. Theirs ranged from  $-.44$  to  $.54$ . In the present male sample the number of children correlated  $-.28$  ( $p < .01$ ) with time to next older, and  $-.19$  ( $p < .05$ ) with average time between births, while the other two *rs* did not reach significance; none of the four *rs* reached significance in the female sample.

These interrelationships are more complicated than appears on the surface. To begin with, the *rs* are not constant: in calculations concerning the time span to next younger all single and youngest Ss fall away; for time span to next older all single and eldest Ss fall away; for the average time span all single Ss fall away; while for *rs* between the time spans to next younger

and next older all single, youngest, and eldest Ss fall away. Intercorrelations are thus not directly comparable.

Furthermore, the abovementioned part-whole  $r$ s vary with family size. When  $S$  is a single child, there are no time span variables to be considered; however, they are not equal to zero, since that is the case only with twins. When there are three children in the family and  $S$  is the eldest, only time span to next younger is considered; when  $S$  is the youngest child, only time span to next older is considered; while when  $S$  is the middle child both time spans enter into the average. When there are four or more children in the family, there is less of a relationship between the two separate time spans and the average time span, since more data enter into the average. Variations of these influences would tend to lower the  $r$  between the average time span variable and the other three constellation variables, since they do not all have an effect in the same direction.

An important difference between the samples of Waldrop and Bell's studies (19, 20) and the present study is likely to have resulted from the age of the Ss. In the present samples of university students the proportion of completed families was probably higher than in their samples of nursery school boys and newborns. Allowance has to be made, for instance, for late-comers in the families of older parents, which would increase the time span variables. Probably this factor, together with the possibility of less family planning, is reflected in the higher means and standard deviations for the time span variables in the present study than in Waldrop and Bell's [e.g., see Waldrop and Bell (20) for average time between births:  $\bar{X} = 44.84$  months,  $SD = 21.44$  for combined male and female sample, compared to 26.54 and 7.79].

All of this leads to the conclusion that the existence of a single family-size-and-density factor cannot be assumed for all samples.

In deriving their index, Waldrop and Bell calculated standard scores for each variable, multiplied each standard score by the appropriate factor loading, and summed over the four components. They remarked that since the differences between factor loadings were small, there would be no great loss of precision in computing the index by simply combining standard scores for the four variables. In the absence of any indication of the relative contribution of the four variables, the latter procedure was followed in the present study.

The  $r$ s between this index and the HT and RAM scores were nonsignificant in both the male and female samples, ranging from .02 to .12.



### 3. Dyads

In the male sample the distribution of the sex/ordinal position dyads was as follows: M1M = 32, MM2 = 32, M1F = 30, and FM2 = 19, leaving 45 Ss in later ordinal positions. In the female sample the distribution was as follows: F1F = 23, FF2 = 28, F1M = 27, and MF2 = 28, leaving 54 in later ordinal positions. One-way analyses of variance between these dyads for the HT and RAM resulted in nonsignificant  $F$  values in both male and female samples.

### 4. Interactions

Two-way analyses of variance showed no significant birth order  $\times$  family size interactions for either HT or RAM in either the male or female samples. Significant interaction between fathers' occupational status (4) and family size was found for RAM in the male sample ( $F = 4.47$ ,  $p < .01$ ), and for HT in the female sample ( $F = 2.35$ ,  $p < .05$ ). A significant occupational status  $\times$  family size and density index interaction was found for HT in the male sample ( $F = 2.23$ ,  $p < .05$ ). The remaining five analyses of variance involving occupational status did not reveal significant interaction effects.

### D. DISCUSSION

In addition to the more traditional comparisons between ordinal positions, the present study used family constellation variables that have not been used in this context before. The negative findings have to be considered from various points of view.

The low reliabilities of the HT and RAM in the present samples accounted, at least in part, for low relationships with other measures. However, the indications of predictive validity in even such small samples should, to some extent, be weighed against this.

Another possibility is that the motivational aspects tapped specifically by the HT and RAM are not subject to family constellation influences. Where indications of a relationship between achievement motivation and family constellation measures have been found [*cf.* Atkinson and Miller in Heckhausen (5); also Rosen (12) and Sampson (15)] projective measures have been used, although two TAT studies (9, 14) yielded negative results too. Wolkon and Levinger's (23) negative findings were on the Edwards Personal Preference Schedule. The real issue is, however, not so much which measures have been used, as the validity of the measures under consideration. Apart from the present validity findings, meaningful rela-



tionships to criteria of achievement have been reported for both the HT (6, 7) and the RAM (10, 22).

The negative findings with respect to Waldrop and Bell's variables, as well as the sex/ordinal position dyads, have to be cross-validated on other samples. In view of the comments on the Waldrop-Bell approach, it should be clear that it needs careful evaluation in this kind of study and with samples differing from those of its originators.

The three interaction effects tend to confirm the likelihood of complex relationships suggested previously (11, 12, 14). However, these three, out of eight analyses of variance involving occupational status, provided neither positive nor negative conclusive evidence. What these findings did underline was the necessity of large samples for this kind of research, to allow simultaneous control on several variables (16).

Notwithstanding the above considerations, one is inclined to view the present findings as generally in line with other negative findings on the relationship between achievement motivation and family constellation variables.

#### REFERENCES

1. ANASTASI, A. *Psychological Testing* (3rd ed.). New York: Macmillan, 1968.
2. BIGNER, J. J. Sibling influence on sex-role preference of young children. *J. Genet. Psychol.*, 1972, **121**, 271-282.
3. CROWNE, D. P., & MARLOWE, D. *The Approval Motive*. New York: Wiley, 1964.
4. HALL, J., & JONES, D. C. Social grading of occupations. *Brit. J. Sociol.*, 1950, **1**, 31-55.
5. HECKHAUSEN, H. *The Anatomy of Achievement Motivation*. New York: Academic Press, 1967.
6. HOLMES, D. S. Conscious self-appraisal of achievement motivation: The self-peer rank method revisited. *J. Consult. & Clin. Psychol.*, 1971, **36**, 23-26.
7. HOLMES, D. S., & TYLER, J. D. Direct versus projective measurement of achievement motivation. *J. Consult. & Clin. Psychol.*, 1968, **32**, 712-717.
8. KAHN, M. H., MAHRER, A. R., & BORNSTEIN, R. Male psychosexual development: Role of sibling sex and ordinal position. *J. Genet. Psychol.*, 1972, **121**, 187-196.
9. KARABENICK, S. A. On the relation between personality and birth order. *Psychol. Rep.*, 1971, **28**, 258.
10. MEHRABIAN, A. Male and female scales of the tendency to achieve. *Educ. & Psychol. Meas.*, 1968, **28**, 493-502.
11. OBERLANDER, M., JENKIN, N., HOULIHAN, K., & JACKSON, J. Family size and birth order as determinants of scholastic aptitude and achievement in a sample of eighth graders. *J. Consult. & Clin. Psychol.*, 1970, **34**, 19-21.
12. ROSEN, B. C. Family structure and achievement motivation. *Amer. Sociolog. Rev.*, 1961, **26**, 574-585.
13. ROSENBERG, B. G., & SUTTON-SMITH, B. Sibling association, family size, and cognitive abilities. *J. Genet. Psychol.*, 1966, **109**, 271-279.
14. ROSENFELD, H. M. Relationships of ordinal position to affiliation and achievement motives: Direction and generality. *J. Personal.*, 1966, **34**, 467-480.

15. SAMPSON, E. E. Birth order, need achievement, and conformity. *J. Abn. & Soc. Psychol.*, 1962, **64**, 155-159.
16. SHAVER, P., FRENCH, J. R. P., & COBB, S. Birth order of medical students and the occupational ambitions of their parents. *Internat. J. Psychol.*, 1970, **5**, 197-207.
17. SUTTON-SMITH, B., & ROSENBERG, B. G. Age changes in the effects of ordinal position on sex-role identification. *J. Genet. Psychol.*, 1965, **107**, 61-73.
18. SUTTON-SMITH, B., ROBERTS, J. M., & ROSENBERG, B. G. Sibling associations and role involvement. *Merrill-Palmer Quart.*, 1964, **10**, 25-38.
19. WALDROP, M. F., & BELL, R. Q. Relation of preschool dependency behavior to family size and density. *Child Devel.*, 1964, **35**, 1187-1195.
20. ———. Effects of family size and density on newborn characteristics. *Amer. J. Orthopsychiat.*, 1966, **36**, 544-550.
21. WARREN, J. R. Birth order and social behavior. *Psychol Bull.*, 1966, **65**, 38-49.
22. WEINER, B., JOHNSON, P. B., & MEHRABIAN, A. Achievement motivation and the recall of incompleting and completed exam questions. *J. Educ. Psychol.*, 1968, **59**, 181-185.
23. WOLKON, G. H., & LEVINGER, G. Birth order and need for achievement. *Psychol Rep.*, 1965, **16**, 73-74.

*Subject Committee of Psychology*  
*University of Port Elizabeth*  
*P. O. Box 1600*  
*Port Elizabeth, South Africa*

## SELF-ACTUALIZATION AND SELF-IDEAL DISCREPANCY\*

*Virginia Commonwealth University*

---

JOHN MAHONEY AND JACK HARTNETT

---

### SUMMARY

The Time Competence Scale from the Personal Orientation Inventory was administered to 28 male and 47 female college students who then described their real and ideal self-concepts using the Interpersonal Check List. Self-actualizing individuals manifested greater congruence between their real and ideal self-concepts than did nonactualizing persons. Dominance was seen as a significantly more desirable integrative mechanism by the nonactualizing subjects. There was a nonsignificant trend in the desirability of an increase in acceptance as a function of the degree of self-actualization. Findings are discussed in terms of individuals' perceptions of congruence with their environment.

---

### A. INTRODUCTION

Maslow (8, p. 197) has stated, "Self-actualization is defined in various ways but a solid core of agreement is perceptible. All definitions accept or imply . . . acceptance and expression of the inner core of self, i.e., actualization of these latent capacities, and potentialities, 'full functioning,' availability of the human and personal essence."

The importance of self-acceptance for full functioning is not unique to Maslow's theory. This is a view shared by Goldstein (4) and Rogers (9). The Rogerian system, developed primarily from clinical observation, also emphasizes the importance of self-acceptance, arguing that a basic congruence between the real self and the ideal self is an absolute prerequisite for full functioning. If the discrepancy between the real self and the ideal self is great, the person will be discontented and maladjusted. It is clear that self-actualization and self-acceptance are not equivalent, but that the latter is a proper subset of the former; there can be no self-actualization without self-acceptance.

Maslow has argued that the characteristics of the self-actualizing person

---

\* Received in the Editorial Office on May 16, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

include the lack of hostility and the acceptance of the self, other persons, and the natural world for what they are. Since the self-actualizing person functions efficiently within the world as it is and does not attempt to overcome problems by physical or psychological dominance, there should be minimal discrepancy for the self-concept between actual and desirable degrees of dominance for the self-actualizing individual. By contrast, the nonactualizing personality should view greater dominance as desirable, since he maintains interpersonal patterns characterized by manipulation and rejects the more symbiotic acceptance-oriented relationships as leaving himself exposed to exploitation by others (11). This leads to a prediction regarding the perception of dominance.

*Hypothesis I.* Nonactualizing persons will perceive a greater discrepancy between actual and desirable magnitude of dominance in interpersonal relationships than will self-actualizing persons.

*Hypothesis II.* To the extent that a discrepancy between the real- and ideal-self exists for the self-actualizing person, it will be in the desire for acceptance, rather than dominance.

## B. METHOD

### 1. Subjects

There were 75 volunteer Ss, 28 male and 47 female, all of whom were enrolled in an introductory psychology course at Virginia Commonwealth University. The ages of the subjects ranged from 18 to 28 years. The difference between the mean ages for the males (22.60 years) and the females (21.72 years) was not significant ( $t = 1.14$ ,  $df = 73$ ,  $p = ns$ ).

### 2. Measures

The Time Competence Scale (TC) from the Personal Orientation Inventory by Shostrom (10), consisting of 23 items, was used as the measure of self-actualization. The time competent person is described as one who lives primarily in the present with full awareness, contact, and full feeling reactivity, while the time incompetent person lives primarily in the past, with guilts, regrets, and resentments, and/or in the future, with idealized goals, plans, expectations, predictions, and fear. The TC Scale has emerged as a general single best estimate of self-actualization in a variety of studies (5, 6). Damm (2) has indicated that the simple raw TC score yields the highest overall correlation between all Personal Orientation Inventory indices

of self-actualization, identifying people who are oriented to the present as contrasted to concern with past guilt or future anxiety. In addition, the psychometric advantages of moderately high reliability and relative brevity of the TC Scale commend it for survey work.

The measures of dominance and acceptance were derived via mathematical formulae from Form IV of the Interpersonal Check List (ICL) developed by Leary (7). The ICL consists of 128 words and phrases, and Foa (3) has demonstrated that the ICL yields two orthogonal factors, termed Dominance (DOM) and Love (LOV), in accordance with the theoretical framework developed by Leary (7). The statistical independence of the two dimensions has also been demonstrated by Bieri and Lobeck (1) and Wiggins (12).

### 3. Procedure

The Time Competence Scale and the Interpersonal Check List were administered in counterbalanced order to control for order effects. Each subject was categorized as either self-actualizing (SA), normal (N), or non-actualizing (NA) on the basis of responses to the Personal Orientation Inventory and the cutoff points provided by the Time Competence Scale.

In addition, the nonactualizing group was further divided to yield two classes of nonactualizing persons, those scoring near the mean for the Time Competence Scale, and those scoring much below the mean. This division yielded two groups of nonactualizing persons, designated as Moderately Nonactualizing (MNA) and Severely Nonactualizing (SNA). The frequencies for each group were as follows: SA, 12; N, 24; MNA, 24; SNA, 15. There was no significant difference between males and females in categorical placement ( $\chi^2 = 2.57$ ,  $df = 3$ ,  $p = ns$ ).

For the present study the ICL was prepared in card form, and each S was directed to sort the cards into two piles, choosing the adjectives and phrases that described "yourself—the way you really are" for one pile, and the remaining cards in the other. After the cards were scored, they were shuffled and returned to the S for a second sorting, this time with instructions to choose the adjectives and phrases that describe your "ideal self—the way you want to be."

By use of the ICL norms provided by Leary (7), standard scores on the love and dominance dimensions were derived for the real self and the ideal self. The difference between the standard scores for the real self and the ideal self was used as the measure of self-acceptance on each of the two dimensions.



TABLE 1

MEAN STANDARD SCORES AND DIFFERENCES FOR REAL AND IDEAL SELF ON DOMINANCE (DOM) AND LOVE (LOV) DIMENSIONS FOR SEVERELY NONACTUALIZING (SNA), MODERATELY NONACTUALIZING (MNA), NORMAL (N), AND SELF-ACTUALIZING (SA) GROUPS

Dimension	SNA	MNA	N	SA
DOM				
Real	56.3	53.8	55.3	60.7
Ideal	64.7	66.5	64.7	62.8
Difference	8.4	12.7	9.4	2.1
LOV				
Real	52.1	51.2	49.7	47.8
Ideal	53.3	54.9	54.7	53.1
Difference	1.2	3.7	5.0	5.3

## C. RESULTS

A preliminary regression analysis was performed to assess the effect of sex differences. No significant difference emerged for sex as component of DOM ( $F = .07$ ,  $df = 1/73$ ,  $p = ns$ ), or LOV ( $F = 1.78$ ,  $df = 1/73$ ,  $p = ns$ ). Accordingly, the data for the two sexes were combined.

The mean standard scores for the real and ideal self on the DOM and LOV dimensions are presented in Table 1, along with the mean ideal-real discrepancy for each dimension for the four groups. The results of analyses of variance for the difference scores for the two dimensions are presented in Table 2. A Neuman-Keuls test demonstrated that the SA group was significantly lower in DOM discrepancy ( $p < .01$ ) than the MNA group and the N and SNA groups ( $p < .05$ ). None of the other comparisons was significant.

An analysis of variance for the difference scores for the LOV dimension indicated that no significant difference between the self-actualizing and the

TABLE 2  
ANALYSES OF VARIANCE FOR DOMINANCE (DOM) AND LOVE (LOV) DISCREPANCY BETWEEN REAL AND IDEAL SELF

Source	SS	df	MS	F	p
DOM					
Between	888.784	3	296.261	3.893	.025
Residual	5402.363	71	76.089		
Total	6291.147	74			
LOV					
Between	165.077	3	55.025	0.685	ns
Residual	5696.443	71	80.231		
Total	5861.520	74			

nonactualizing groups was present, although inspection of the means in Table 1 reveals that a trend exists in the hypothesized direction.

#### D. DISCUSSION

The results tend to support the notion that differences between self-actualizers and nonactualizers in terms of self-acceptance is dependent on the specific traits that are measured. Inspection of the scores on the dominance dimension shows that while the self-actualizing group had a higher real-self perception score than the other groups, they had a lower ideal-self perception score than nonactualizers. On the other hand, the love dimension scores show a different trend, with self-actualizers having a lower real-self score than the other groups and a lower ideal-self score than nonactualizers although differences in the ideal-self are slight.

The finding that self-actualizers perceived themselves as more dominant and less loving than the nonactualizers saw themselves is interesting in light of the fact that a self-actualizing person is typically considered to be an individual who is more loving and less concerned with the domination of others than the average person. One interpretation of the findings is that the self-actualizers were less concerned with giving a socially desirable response and hence felt few qualms about describing themselves as more dominant and less loving than others. Another plausible explanation for the findings is that self-actualizers were more accurate in their perceptions of themselves and set a realistic ideal-self goal in line with those perceptions.

A final interpretation can be made in light of the possibility that a self-actualizing person possesses sufficient dominance in interpersonal relations to fulfill his specific needs and thus remain self-actualized. A nonactualizing type, however, seeks dominance as a possible alleviation for his inherent feelings of frustration that arise from his inability to fulfill his creative potential. Both of these life styles are self-sustaining, however. An actualizing person continues to actualize because he has found the correct mechanism for symbiosis. A nonactualizing person seeking dominance as a solution to his frustrations cannot fulfill his being-needs effectively, since he is oriented toward dominance of the social environment. Thus the nonactualizing person is living an essentially predatory life style. He sees himself as having to work to achieve actualization, rather than existing as a harmoniously integrated element of society.

#### REFERENCES

1. BIERI, J., & LOBECK, R. Self-concept differences in relation to identification, religion and social class. *J. Abn. & Soc. Psychol.*, 1961, **62**, 94-98.

2. DAMM, V. J. The relationship of ego-strength of creativity and intelligence in high school students. Unpublished Doctoral dissertation, University of Oregon, Eugene, Oregon, 1967.
3. Foa, U. G. Convergences in the analysis of the structure of interpersonal behavior. *Psychol. Rep.*, 1961, **68**, 341-353.
4. GOLDSTEIN, K. *The Organism*. New York: American Book, 1939.
5. KNAPP, R. R. Relationship of a measure of self-actualization to neuroticism and extraversion. *J. Consult & Clin. Psychol.*, 1965, **29**, 168-172.
6. LABACHE, P. Self-actualization in college students: Interrelationships of self-actualization, personal characteristics and attitudes in subcultures of liberal arts freshmen and seniors. Unpublished Doctoral dissertation, Appalachian State University, Boone, North Carolina, 1970.
7. LEARY, T. *Interpersonal Diagnosis of Personality*. New York: Ronald, 1957.
8. MARLOW, A. H. *Toward a Psychology of Being* (2nd ed.). New York: Van Nostrand Reinhold, 1968.
9. ROGERS, C. R. *On Becoming a Person*. Boston: Houghton Mifflin, 1961.
10. SHOSTROM, E. L. *Manual for the Personal Orientation Inventory*. San Diego: Educ & Indust Test. Serv., 1966.
11. ———. *Man, the Manipulator*. New York: Abingdon Press, 1967.
12. WOODING, N. A. Structural aspects of the interpersonal system of personality diagnosis: Level II. Master's thesis, Stanford University, Stanford, California, 1961.

*Psychology Department*

*Virginia Commonwealth University*

*810 West Franklin Street*

*Richmond, Virginia 23220*

## THE CONTROL ORIENTATION OF BLACK ATHLETES IN RELATION TO UNOBTRUSIVE TASKS OF SKILL\*

Boise State College

STEVEN THURBER, DEL HEACOCK, AND BOOKER BROWN

### SUMMARY

Differential predictions have been made concerning the effect of external beliefs on the skilled performance of Negroes. In the present study it was found that among black college athletes, externality covaried with elevated attainment on unobtrusive tasks.

### A. INTRODUCTION

Past research has shown that Negroes tend to perform less adequately than whites on laboratory and academic tasks requiring skill (2, 4). Blacks have also been found to be external in locus of control when compared to whites (1). This suggests that Negroes may have low expectancies with regard to the self-determination of reinforcements in achievement situations. It has therefore been hypothesized that the external beliefs of Negroes may function to attenuate skilled performance (5).

In contrast, Gurin *et al.* (3) have argued that the negative experiences of the Negro (e.g., discrimination, segregation) have fostered the realistic growth of external beliefs. They further state that externality may actually be an adaptive and functional orientation for Negroes.

### B. METHOD

Ten athletes (five black, five white), comprising the varsity basketball team at an institution classified as a major college in terms of intercollegiate participation, were administered the I-E scale of Rotter (6). Their scores were then correlated with official team statistics compiled after 19 games. The data considered important for the present analysis were total points scored and assists (passes leading to baskets).

### C. RESULTS

As expected, blacks ( $\bar{X} = 13.8$ ) were significantly more external than whites ( $\bar{X} = 9.6$ );  $t = 3.13$ ,  $df = 8$ ,  $p < .01$ . A correlation of  $+ .95$  ( $p <$

\* Received in the Editorial Office on May 16, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

.01) was obtained for black players between locus of control and total points scored. The corresponding relationship for white players was  $-.395$ , which was not significant. Consistent but nonsignificant correlations of  $+.56$  and  $-.496$  were obtained for black and white players, respectively, between locus of control and number of assists.

In comparison to previous research, the present study was unique in that it investigated locus of control correlates among highly skilled performers in a domain in which Negroes traditionally excel. Although the  $N$  was unavoidably small, the study did have the advantage of what might be considered relatively stable and unobtrusive outcome measures. The findings indicate support for the notion that externality in blacks may be associated with enhanced performance on skill-demanding tasks.

#### REFERENCES

1. BATTLE, E. S., & ROTTER, J. B. Children's feelings of personal control as related to social class and ethnic group. *J. Personal.*, 1963, **31**, 482-490.
2. DREGER, R. M., & MILLER, K. S. Comparative psychological studies of Negroes and whites in the United States. *Psychol. Bull.*, 1960, **57**, 361-402.
3. GURIN, P., GURIN, G., LAO, R., & BEATTIE, M. Internal-external control in the motivational dynamics of Negro youth. *J. Soc. Issues*, 1969, **25** (3), 29-53.
4. KATZ, I., GOLDSTON, J., & BENJAMIN, L. Behavior and productivity in bi-racial work groups. *Hum. Relat.*, 1958, **11**, 123-141.
5. LEFCOURT, H. M., & LADWIG, G. W. The American Negro: A problem in expectations. *J. Personal. & Soc. Psychol.*, 1965, **1**, 377-380.
6. ROTTER, J., SEEMAN, M., & LIVERANT, S. Internal versus external control of reinforcement: A major variable in behavior theory. In N. F. Washburn (Ed.), *Decisions, Values, and Groups* (Vol. 2). London: Pergamon Press, 1962. Pp. 473-516.

*Department of Psychology*  
*Boise State College*  
*1907 Campus Drive*  
*Boise, Idaho 83707*



## THE RELATIONSHIP OF MARIHUANA USAGE TO PERSONALITY AND MOTIVATIONAL FACTORS\*

*Wichita State University*

---

CHARLES BURDSAL, GARY GREENBERG, AND RANDIE TIMPE

---

### SUMMARY

The 16 Personality Factors Questionnaire (16 PF), Motivation Analysis Test (MAT), and a marihuana usage questionnaire were administered anonymously to 104 undergraduate students. Raw scores were converted to sten scores to eliminate sex and age differences. Product-moment correlations were computed on data from the questionnaires. An iterative principal axis solution was applied to the correlation matrix followed by Kaiser Vari-max orthogonal rotation and graphical oblique rotations.

The most significant finding was that marihuana users were not a homogeneous group in terms of personality and motivational structure. Four identifiable personality and motivational patterns were found to be related to such use: (a) an antisocial norm group; (b) a frustrated upper-middle class group; (c) a hostile rebel group; (d) a follower group. None of these indicate pathological patterns.

---

### A. INTRODUCTION

In recent years both the public and the scientific communities have over-reacted to the almost overwhelming increase in the use of psychedelic drugs. Indeed, the "marihuana problem" (11) has led to a government study of drug use (12). Clearly, the time is ripe for controlled investigation of drug use and users.

The voluminous amount of research on alcoholism and hard drug use suggests that a major step in understanding the "psychedelic drug problem" is to identify the personality traits of the users. The assumption is that identification of the potential user would lead to the eventual control of drug use. Indeed, several recent studies have taken this very approach.

Thus, Spevack, Pihl, and Sternthal (14) sent questionnaires to a sample of 400 high school students in an attempt to identify the "motivational"

---

\* Received in the Editorial Office on May 17, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

factors that predisposed the users in the sample (13.2% of the 260 respondents) to use regularly or experiment infrequently with psychedelic drugs. These investigators reported that conformity to teenage subculture norms was *the* primary motive for drug use in their sample.

Similarly, Keeler (9) questioned 54 marihuana users about the motives involved in their use of the drug. Twenty-eight of his Ss suggested that curiosity was principal in their use of marihuana; only six Ss reported their primary motivation for use to be conformity; 12 users gave equal weight to the two reasons.

Other studies have sought to identify personality characteristics among drug users, the assumption being that a particular "type" of personality is involved. Thus, Brill, Crumpton, and Grayson (2) administered several Minnesota Multiphasic Personality Inventory scales to a group of college students whose marihuana usage varied from "frequent use" to "never use." They found that the frequent user of marihuana was more hostile or rebellious and tended to seek more stimulation than the nonuser. In addition, they indicated that as a group, the frequent marihuana users reported more long-standing emotional problems than the nonusers. Another study of personality factors in psychedelic drug users (10, p. 70) reported that "Drug users presented themselves as being less acculturated and more non-conformist than the nonusers, manifesting more anxiety and less effective superego controls, with greater potential for asocial rather than antisocial acting out."

Several comments about this type of research are in order: First, a search of the literature reveals that there is a limited amount of this research. Surely, more is called for. Second, the few studies that do exist seem to suggest that a particular personality type is involved in psychedelic drug use. For example, Klechner (10, p. 66) stated that "It is felt that there may be a specific personality factor constellation, syndrome, or type characteristic of the psychedelic drug user . . ." Assertions of this type seem to be unwarranted in light of what we know about behavior. Also, when we consider the fact that 40% of all 18 to 25 year olds have at least tried marihuana (12), it is unlikely that a group this size is homogeneous in personality structure. Indeed, as early as 1953 Becker (1) pointed out that the presence of a given kind of behavior (e.g., drug use) is a function of a long complex of social experiences during which the individual comes to acquire a meaning of the behavior. In addition, this social learning process provides the individual with perceptions and judgments of objects and sensations, making possible and desirable (or reinforcing) a

particular activity (in this present case, drug use). Rather, then, than suggesting that one personality type is predisposed to drug usage, it seems more likely that any personality type can acquire drug use habits. Third, the studies in this area typically assume that drug use and pathological behavior are intimately connected. Hence the use of such descriptive terms as "oral dependent type," "less acculturated," "asocial" (10), "oedipal revenge" (7), and "amotivational syndrome" (11). However, current notions about pathological behavior seem to indicate that everyone is pathological to a degree, and it may be virtually impossible to decide between the sane and the insane (13, 15). If this is the case, it hardly appears fruitful to search for a pathological background in the drug user.

Finally, the particular questionnaires and personality tests used in these studies often have not been the most appropriate for such purposes. Such tests may have been hastily constructed, poorly used, and their results inadequately analyzed. Thus, the present study was designed to determine (a) if a personality type is involved in psychedelic drug use, by (b) using tests whose reliability has been clearly established [Cattell's (5) 16 Personality Factors Questionnaire (16 PF) and Cattell's (3) Motivation Analysis Test (MAT)] and (c) analyzing the results of these tests with factor analytic techniques. Whereas at least one investigator did use the Cattell 16 PF (10), his results were analyzed by means of many multiple Fisher *t*-test comparisons. This procedure is questionable at best. However, a factor analytic approach has the advantage of being able to pick out related personality dynamics among the various scales within the tests employed.

## B. METHOD

Cattell's 16 PF (5) and his MAT (3), along with a marihuana usage questionnaire, were administered to 104 undergraduate students in psychology classes at Wichita State University. It should be noted that one of the courses was a night class which was more representative of the general public than the usual psychology class. All test results were collected on an anonymous basis. The mean age of the students was 24.4 years. By conversion of raw scores obtained from the tests to sten scores, differences in the personality and motivational factors resulting from age and sex variations were eliminated.

The marihuana usage questionnaire involved an eight-point self-rating scale, scored zero to 7. This scale was adapted from the National Commission on Marihuana and Drug Abuse (12) report on marihuana, and involved

the subjects selecting the category into which they best seemed to fit. The categories ranged from "having never used marihuana" and "having used it at least once but not now using it" on the low end to "currently using it once per day" and "currently using it several times per day" on the high end.

Product-moment correlations were calculated among the scores obtained from the questionnaires proceeding to a factor analysis described in the following section.

### C. RESULTS<sup>1</sup>

Eigenvalues were computed for the correlation matrix. The Scree test (4) indicated 15 factors. An iterative principal axis solution was applied to the correlation matrix until the communalities stabilized in the third decimal place. A Kaiser Varimax Orthogonal Rotation (8) was applied to the factor matrix followed by five graphical oblique rotations (6), resulting in a 73 percent .10 width hyperplane. Four factors were found to be related to marihuana usage. The salient loadings (absolute value greater than .2) for these factors may be found in Table 1.

### D. DISCUSSION

Four distinct personality-motivational patterns were found to be related to marihuana usage. Although, with the exception of Factor IV, the results demonstrated little relation to the usual 16 PF second order factors, they did appear to represent distinct personality types.

Factor I represents an antisocial norm group. Persons high on this factor appear to have rejected the work ethic, as indicated by a low integrated (I) Ca score on the MAT, and are not motivated by fear (MAT I Fr—); *i.e.*, they are unafraid of being caught and unlikely to worry about the future in terms of health, insurance, etc. They also seem unable to deal with direct confrontation (MAT I Pg—). This group tends to think in concrete terms (16 PF B—) and has some difficulty in making logical generalizations from concrete situations at hand. Finally, these individuals have a high sensitivity to emotional and physical pain (16 PF I+), as well as a tendency to be overprotected.

The examination of the motivational scales that loaded on the second factor related to marihuana use suggests that Factor II represents a frustrated upper-middle class group. The very low unintegrated (U) motiva-

<sup>1</sup> For such data as the intercorrelations,  $V_o$ ,  $V_{rs}$ ,  $R_f$ , etc., and the marihuana questionnaire, write Charles Burdsal at the address shown at the end of this article.



TABLE 1  
SALIENT LOADINGS OF THE FACTORS RELATED TO MARIHUANA USAGE

Measure	Subscale	Loading
<i>Factor I</i>		
MAT	Integrated Ca (Career)	-.543
MAT	Integrated Fr (Fear)	-.451
MAT	Integrated Pg (Pugnacity)	-.314
16 PF	B (Concrete vs. Abstract Thinking)	-.387
16 PF	I (Tough-Minded vs. Tender-Minded)	.340
Marihuana usage		.533
<i>Factor II</i>		
MAT	Unintegrated Ca (Career)	-.689
MAT	Unintegrated Na (Narcissism)	.595
MAT	Unintegrated Ma (Mating)	.242
16 PF	A (Aloof vs. Warmhearted)	-.238
16 PF	L (Trusting vs. Suspicious)	.250
16 PF	Q <sub>2</sub> (Group Dependent vs. Self-Sufficient)	.441
Marihuana usage		.515
<i>Factor III</i>		
MAT	Unintegrated Pg (Pugnacity)	.349
MAT	Unintegrated As (Assertion)	.412
MAT	Unintegrated Sw (Sweetheart-Spouse)	.238
MAT	Integrated Se (Superego)	-.210
16 PF	F (Serious vs. Enthusiastic)	.600
16 PF	H (Restrained vs. Venturesome)	.243
16 PF	N (Forthright vs. Shrewd)	-.380
Marihuana usage		.507
<i>Factor IV</i>		
MAT	Integrated Ho (Home-Parental)	.219
MAT	Integrated Se (Superego)	-.471
16 PF	B (Concrete vs. Abstract Thinking)	-.431
16 PF	E (Humble vs. Assertive)	-.801
16 PF	H (Restrained vs. Venturesome)	-.331
16 PF	L (Trusting vs. Suspicious)	-.251
Marihuana usage		.323

Note: MAT = Motivation Analysis Test; 16 PF = 16 Personality Factors Questionnaire.

tion in the career area (MAT U Ca—), coupled with high frustration in narcissistic comfort needs (MAT U Na+) and moderately high frustrated sexual needs (MAT U Ma+), describes a group that has rejected the work ethic while either denying or not fulfilling such needs as the desire for the "good things of life" (color television, sports cars, etc.) and desire for sexual activity. This group also tends to be reserved and detached, indicating an objectivity and freedom from social needs (16 PF A—). These people are somewhat jealous and insecure concerning how they are per-



ceived by others (16 PF L+). Finally, they tend to carry out activities themselves rather than seeking help (16 PF Q<sub>2</sub>+). This independence may be a result of a timidity and reluctance to ask others for aid.

The third factor related to marihuana use appears to represent a hostile rebel personality and motivational pattern. These people tend to be hostile (MAT U Pg+) and have unfulfilled status needs (MAT U As+). They also have some unmet need to be loved (MAT U Sw+) and tend to have a moderately low operating superego (MAT I Se—). This group is exhibitionistic, generally seeks to be the center of conversations, and is quite impulsive (16 PF F+). A moderate amount of brashness and social confidence was also found (16 PF H+). Finally, this group is unpretentious and open in their interaction with others, and members sometimes expose themselves to criticism because they share negative as well as positive aspects of their thoughts and behaviors (16 PF N—).

The final factor associated with marihuana use resembles the 16 PF second order factor of follower *vs.* leader. The motivational variables indicate moderate dependency needs (MAT I Ho+) and a rather low operating superego (MAT I Se—). They tend to be concrete in their thinking (16 PF B—), extremely mild and unconcerned with dominating others (16 PF E—), and somewhat shy, timid, and a bit anxious in the social setting. Finally, they tend to be a bit gullible, easy to get along with, and trusting of others (16 PF L—).

### E. CONCLUSIONS

The most significant finding of this study was that the marihuana users in this sample of college students were not a homogeneous group in terms of personality and motivational structure. Four identifiable personality and motivational patterns were found to be related to such use: (a) an anti-social norm group; (b) a frustrated upper-middle class group; (c) a hostile rebel group; (d) a follower group. These dynamics are not conclusive because of a limited sample size (104 students) and may be further clarified by more research utilizing more measures of personality and motivation.

Another conclusion of this research is that although the personality-motivational patterns found might not be considered "the model of mental health," nothing pathological in the usual sense was indicated. This may be due in part to a lack of emphasis on pathological measures, and further research will be needed to demonstrate relationships of pathological personality-motivational patterns to marihuana usage.

## REFERENCES

1. BECKER, H. S. Becoming a marihuana user. *Amer. J. Sociol.*, 1953, **59**, 235-242.
2. BRILL, N. Q., CRUMPTON, E., & GRAYSON, H. M. Personality factors in marihuana use. *Arch. Gen. Psychiat.*, 1971, **24**, 163-165.
3. CATTELL, R. B. Motivation Analysis Test. Institute for Personality and Abilities Testing, Champaign, Illinois, 1964.
4. ———, *Ed.* Handbook of Multivariate Experimental Psychology. Chicago: Rand McNally, 1966.
5. ———. 16 Personality Factors Questionnaire. Institute for Personality and Abilities Testing, Champaign, Illinois, 1967.
6. CATTELL, R. B., & FOSTER, M. J. The rotoplot program for multiple single-plane, visually guided rotation. *Behav. Sci.*, 1963, **8**, 156-165.
7. FREEDMAN, D. X. The use and abuse of psychedelic drugs. *Bull. Atomic Sci.*, 1968, **24**(4), 6-14.
8. HARMAN, H. H. Modern Factor Analysis. Chicago: Univ. Chicago Press, 1967.
9. KEELER, M. H. Marihuana induced hallucinations. *Dis. Nerv. Syst.*, 1968, **29**, 314-315.
10. KLECHNER, J. H. Personality differences between psychedelic drug users and non-users. *Psychology*, 1968, **5**, 66-71.
11. MCGLOTHLIN, W. H., & WEST, L. J. The marihuana problem: An overview. *Amer. J. Psychiat.*, 1968, **125**, 370-378.
12. NATIONAL COMMISSION ON MARIHUANA AND DRUG ABUSE. Marihuana: A Signal of Misunderstanding. New York: Signet-New Amer. Lib., 1972.
13. ROSENHAN, D. L. On being sane in insane places. *Science*, 1973, **179**, 250-258.
14. SPEVACK, M. G., PIHL, R. D., & STERNTHAL, B. A survey study on the motives for drug use by adolescents. Paper read at the Canadian Psychological Association Meeting, Winnipeg, Manitoba, Canada, May 1970.
15. SZASZ, T. The Manufacture of Madness: A Comparative Study of the Inquisition and the Mental Health Movement. New York: Harper & Row, 1970.

*Department of Psychology*  
*Wichita State University*  
*Wichita, Kansas 67208*



## HUMOR JUDGMENTS AS A FUNCTION OF IDENTIFICATION CLASSES: CANADIAN VS. AMERICAN\*<sup>1</sup>

University of Windsor, Canada

LAWRENCE LA FAVE, KEVIN MCCARTHY, AND JAY HADDAD

### SUMMARY

The present experiment tests Hobbes' superiority theory of humor in relation to *identification classes* (a replacement for the construct of reference group)—i.e., a *vicarious superiority* theory of humor is examined. Joke content concerns Canadian-American relations. Ss were selected so as to be either pro-Canadian Canadians or pro-American Americans. Consistent with prediction ( $p < .01$ ) each of the two groups tended to find that permutation of jokes funnier in which its positive identification class was victorious and its negative the butt than was the opposite permutation in which the negative identification class was victorious and the positive the butt.

### A. INTRODUCTION

Hobbes' *superiority theory* of humor has been tested several times. The first two attempts (8, 15) were equivocal in their findings. However, their failures seemed to the authors themselves, in serendipitous hindsight, the consequences of inadequate methodologies, rather than because Hobbes was wrong, as La Fave (4, 5) has indicated.

The first successful attempt appears to have been by La Fave (3). Priest (9) believes he has supported superiority theory also, and despite a methodological problem pointed out by authors Priest and Abrahams (10), their results do also seem to them consistent with the superiority assumption. So also seem to be experiments by La Fave, Haddad, and Marshall (7) and Zillmann and Cantor (16). More particularly, these experiments apparently support a *vicarious* superiority theory, since the Ss making the humor judgments were never themselves either butt or victor in a given joke.

\* Received in the Editorial Office on May 21, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> We would like to thank the following University of Windsor students for their assistance in carrying out this experiment: Susan Viveash, Douglas Baptiste, Susan Robinet, Cathy Williams, and Joseph Pekar.

Rather, Ss either vicariously identified with such butts or victors, or else detached themselves, or negatively related to these butts or victors.

Although the construct *reference group* has proven very useful in making the necessary transition from superiority to *vicarious* superiority humor theory, the reference group construct is not without its problems. For this reason the present paper substitutes the construct of identification classes (ICs). La Fave (5, pp. 208-209) has shown eight ways in which the ICs construct improves upon the reference group. However, the present experiment is relevant to reference-group theory also, at least when the ambiguous word *reference* is understood to mean *identification*, rather than *comparison* (2) or *perspective* (13).

All experiments other than those performed by the La Fave group (5) seem to have suffered from technical inadequacies in attempting to test vicarious superiority humor hypotheses. Specifically, among the procedural errors those other experiments have typically committed are one or more of the following: (a) confounding membership and reference group; (b) failing to control the heard-before variable; (c) failing to employ each S as his own control; (d) misusing statistics, such as employment of analysis of variance or covariance when neither *homoscedasticity* nor more than ordinal scalability could reasonably be assumed (e.g., 4; 5, p. 206). Criticism (d) seems appropriate to the Zillmann and Cantor (16) experiment.

The major reason technical problems loom so large in testing vicarious superiority humor theory is that conventional theory-testing experimental procedure seems inappropriate. By such conventional experimental procedure we mean essentially a procedure which, stripped to its bare essentials, selects random samples from the *same* population, treats these samples *differently*, then predicts *different* results. Such conventional experimental procedure seems typically inappropriate in our research for two basic reasons: (a) Our humor-judgment experiments represent *disguised* attitude measures. More specifically, if a vicarious superiority humor hypothesis is substantiated, one can then use that hypothesis as a premise. Thus, if S finds jokes funnier when IC A triumphs over IC B than conversely, then S apparently prefers A to B. Conventional experimental method is likely to be *invalid* when disguised attitude measures are needed, as Ss are too likely to "see through" the purpose of the experiment, since it must employ a before-after attitude change design or some other similarly transparent tactic. (b) Measurement of vicarious superiority humor effects is too *imprecise* at this stage to expect attitudes of relatively weak ego-involvement to reveal themselves. However, if two samples are randomly



selected from the same population, only relatively weak attitudinal differences can usually be built into the experimental and control groups. (Both the above limitations of conventional experimental procedure could be overcome if *E* had *total* control over *Ss'* environment for a *long* time, which *Es* in this area to date simply have not had.) For these two reasons we follow the Sherif-Hovland social-judgment approach (6, 11, 12). This approach chooses *Ss* from *different* populations, treats them *similarly*, then predicts *differences* between groups.

The Sherif-Hovland procedure permits ego-involving attitudinal differences to be measured and also more readily disguises *E's* intent from *Ss*. However, their approach pays a heavy price in another area. Can one call such a procedure an experiment? Or does one have a mere matching design (since one cannot be certain that no other important differences exist between the groups besides the independent variable)? To offset the above criticism that we would have a mere matching design, rather than a controlled experiment, three basic innovations are introduced:

First, as Sherif, Sherif, and Nebergall (11) have shown, the procedure can be applied not just once, but several times, across a variety of social issues. If the hypothesis persists in being substantiated across such a variety of social issues, then the alternative interpretation (that some uncontrolled variable other than the independent one explains the results) becomes less plausible. A second innovation would be to use each *S* as his own control. Siegel observes as follows (14, p. 62):

A matching design is only as good as the experimenter's ability to determine how to match the pairs, and this ability is frequently very limited. This problem is circumvented when each subject is used as his own control; no more precise matching is possible than that achieved by identity.

However, Siegel is a bit too sanguine on the virtues of using each *S* as his own control, at least in this area. For instance, suppose jokes that one has heard before are thereby perceived as less funny than those one has not. Imagine also that jokes that make one's own group the butt are more likely to have "made the rounds" in that group than in other groups. Now suppose we found, consistent with vicarious superiority humor prediction, that one permutation of jokes (anti-American-pro-Canadian) tends to be funnier to Canadians than to Americans, whereas the other permutation (anti-Canadian-pro-American) tends to be funnier to Americans than to Canadians. Then, despite the fact we had used each *S* as his own control by employing both possible permutations, the "prediction" could prove right for the wrong reason (i.e., failure to control the heard-before variable). Our

third innovation then was to eliminate heard-before jokes. La Fave (3) had combined several techniques to accomplish this: (a) Construct as many original jokes as possible. (b) Eliminate well-known jokes. (c) Revise unoriginal jokes to make them sound unfamiliar. (d) Eliminate any unoriginal jokes placed in the heard-before category by more than one *S* in a pretest. The present experiment employed the first three of the above techniques.

ICs are attitudes. Thus, consistent with the definition of an attitude we prefer, any IC has both an emotive and a cognitive component. The cognitive component of an IC is a *set* of elements. The emotive component may vary on a continuum from hate (highly negative) through neutral to love (highly positive). The *S*'s set of ICs represents a proper subset of his set of attitudes. Further, this paper is only concerned with that proper subset of *S*'s ICs in which the elements refer to *persons*. Also, *S* possesses a given IC if and only if *S* has some symbol (such as a noun or noun phrase) by which he represents this IC to himself.

The hypothesis of the present experiment is as follows: *S* will judge funnier that joke stimulus that esteems his positive IC and disparages his negative IC than a joke stimulus that disparages a positive IC and esteems a negative.

## B. METHOD

### 1. Subjects

A total of 71 *Ss* were selected so that 30 were believed to be pro-Canadian Canadians and the other 41 pro-American Americans. All of the latter were students at either a public university in Detroit (Wayne State) or a private university (University of Detroit). The former were all students at the University of Windsor. Each *S* was categorized as pro-American or pro-Canadian if and only if *S* satisfied two criteria: First, *S* needed to be classified in one of these two classes by a participant observer (i.e., "spy"). After judging the jokes, *S* needed to answer the following question consistent with the category he had been placed into by the participant observer. "Check which country you *prefer* to consider yourself as belonging to: Canada—United States — Other —."

The data of any *S* whose answer on the questionnaire was inconsistent with the participant observer's expectation were (consistent with procedure decided upon *before* the experiment) excluded from the study.

## 2. Procedure

*a. Construction of joke stimuli.* Twenty jokes were constructed, all concerning Canadian-American relations. Every joke involved dialogue between a Canadian and an American. In one permutation ( $P_1$ ) of 10 jokes the Canadian was always victorious and the American invariably the butt of the joke. In the other permutation ( $P_2$ ) of 10 jokes the Canadian was the butt and the American invariably victorious.

*b. Instructions to Ss.* Ss were tested in groups of five or more and told they need *not* write their names. (These measures were taken to insure Ss anonymity—to avoid invalid results as a consequence of trying to please *E*.) Ss in any group tested were (on the basis of information from informants as outlined above) all categorized prior to the experimental session as either pro-Canadian or pro-American. Thus individual names were not needed to test the hypothesis.

Ss were requested to sort each of the 20 jokes under one of five appropriate label cards: "Very Funny"; "Funny"; "Indifferent"; "Unfunny"; and "Very Unfunny." After doing so, they were instructed to complete a very brief questionnaire—asking for age, sex, nationality, last year of school completed, and country of which they preferred to consider themselves members. These questionnaire data needed be consistent with knowledge *Es* obtained about Ss prior to testing. Hence nine of the 50 prospective American Ss were eliminated—seven for failure to identify themselves as Americans, one for returning no questionnaire, and one as a result of having erroneously received the same joke twice.

## C. RESULTS

The five categories into which Ss sorted the jokes were collapsed into two (i.e., a none-some scale) for purposes of statistical analysis—Funny (which included the "Very Funny" and "Funny" labels), and Nonfunny (for the union of the other three categories). Each joke was then scored as to whether it "came out" in the predicted direction. (E.g., a joke from one permutation class,  $P_1$ , was scored consistent with our hypothesis if and only if a greater percentage of pro-Canadian Ss found it funny relative to their mean percentage funniness for all jokes than did pro-American Ss relative to their funniness percentage mean for all jokes.  $P_2$  needed the opposite result. The means of the two groups for all 20 jokes were thus taken into account. Hence, pro-Canadian Ss were expected to be above their mean percentage for  $P_1$  jokes and below for  $P_2$ , while pro-American Ss should be below their mean

percentage for  $P_1$  and above for  $P_2$  jokes.) However, as was also the case for our three previous experiments in this area, only one prediction is really made per joke for a given hypothesis; i.e., the victorious group on a particular joke must be further above its average funniness rating percentagewise for all jokes than is the butt group above its average for all jokes. If  $W$  represents the percent of Ss in the victorious group who found the joke in question funny,  $\bar{W}$  the average funniness which that group assigned to all jokes, and  $L$  and  $\bar{L}$  the results for the losing group on the joke under consideration, then the prediction for a given joke is that  $(W - \bar{W}) - (L - \bar{L}) = +.^2$

Of the 20 jokes 16 scored in the predicted direction. On a one-tailed, one  $df$  test,  $p < .01$ ;  $\chi^2 = 6.8$ .

#### D. DISCUSSION

Although the hypothesis was supported, four of 20 jokes contradicted prediction. (None had in our two previous experiments.) All  $E$ s were situated at the Canadian university, and it seemed obvious these Canadian  $E$ s had not chosen their American Ss as carefully for patriotism as they had their Canadian. We suspected then that it was essentially the American Ss' results that prompted four jokes to contradict prediction. Canadian Ss were below their mean funniness for all jokes on every one of the 10 jokes in which the Canadian was the butt, and were above their mean on nine of 10 jokes in which the Canadian was victorious. However, the American Ss were below their overall mean funniness only four of 10 times when the American lost, and were above their mean only three of 10 times when the American won. It can *neither* be concluded from these data that Canadian Ss went counter to prediction in one of 20 jokes *nor* that American Ss went counter 13 of 20 times. (Such a conclusion is unwarranted because it fails to allow that some jokes are funnier than others.) However, the difference in number of "wrong" Canadian and American predictions (one *vs.* 13) does appear to offer a meaningful comparison:  $\chi^2 = 10.3$ , two-tailed, 1  $df$ ,  $p < .004$ . The best judgment as to why four jokes scored counter to prediction seems to be that American Ss were carelessly chosen.

<sup>2</sup> Parametric statistics seem unwarranted in this experiment because no basis exists to believe scalability is more powerful than ordinal. However, parametric statistics appear unnecessary, since power or efficiency lost was restored by *increasing* the number of Ss. Nor does our hypothesis require a parametric test. Also, since our nonparametric test makes fewer assumptions than a parametric would, by the logician's *revised law of inverse variation* (1, pp. 111-112), a nonparametric test would permit greater theoretic generalization from a substantiated hypothesis.



## REFERENCES

1. COPI, I. M. *Introduction to logic*. New York: Macmillan, 1961.
2. KELLEY, H. H. Two functions of reference groups. In H. Hyman & E. Singer (Eds.), *Readings in Reference Group Theory and Research*. New York: Free Press, 1968.
3. LA FAVE, L. Humor judgments as a function of reference groups: An experimental study. *Diss. Abst.*, 1961, **22**, 1284.
4. ———. Comment on Priest's article: "Election jokes: The effects of reference group membership." *Psychol. Rep.*, 1967, **20**, 305-306.
5. ———. Humor judgments as a function of reference groups and identification classes. In J. H. Goldstein & P. E. McGhee (Eds.), *The Psychology of Humor*. New York: Academic Press, 1972.
6. LA FAVE, L., & SHERIF, M. Reference scale and placement of items with the own categories technique. *J. Soc. Psychol.*, 1968, **76**, 75-82.
7. LA FAVE, L., HADDAD, J., & MARSHALL, N. Humor judgments as a function of identification classes. *Sociol. & Soc. Res.*, in press.
8. MIDDLETON, R. Negro and white reactions to racial humor. *Sociometry*, 1959, **22**, 175-183.
9. PRIEST, R. F. Election jokes: The effects of reference group membership. *Psychol. Rep.*, 1966, **18**, 600-602.
10. PRIEST, R. F., & ABRAHAM, J. Changing reference groups and 1968 election humor. Symposium paper read at Western Psychological Association annual convention, Vancouver, British Columbia, Canada, June, 1969.
11. SHERIF, C. W., SHERIF, M., & NEBERGALL, R. E. *Attitude and Attitude Change: The Social Judgment-Involvement Approach*. Philadelphia: Saunders, 1965.
12. SHERIF, M., & HOVLAND, C. I. *Social Judgment: Assimilation and Contrast Effects in Communication and Attitude Change*. New Haven, Conn.: Yale Univ. Press, 1961.
13. SHIBUTANI, T. Reference groups as perspectives. In H. Hyman & E. Singer (Eds.), *Readings in Reference Group Theory and Research*. New York: Free Press, 1968.
14. SIEGEL, S. *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill, 1956.
15. WOLFF, H. A., SMITH, C. E., & MURRAY, H. A. The psychology of humor: A study of the responses to race disparagement jokes. *J. Abn. & Soc. Psychol.*, 1934, **28**, 341-365.
16. ZILLMANN, D., & CANTOR, J. R. Directionality of transitory dominance as a communication variable affecting humor appreciation. *J. Personal. & Soc. Psychol.*, 1972, **24**(2), 191-198.

*Department of Psychology*

*University of Windsor*

*Windsor, Ontario, Canada N9B 3P4*





## BIRTH ORDER AND RELATED VARIABLES IN A LARGE OUTPATIENT POPULATION\*

*Postgraduate Center for Mental Health, New York City*

---

BERNARD F. RIESS<sup>1</sup> AND JEANNE SAFER

---

### SUMMARY

In an exploratory study of birth order variables in psychiatric patients, 2474 terminated cases at the Postgraduate Center for Mental Health between 1947 and 1962 were collected, and data were obtained on sex, birth order in family, diagnosis, age of sibs, etc. Frequency of ordinal positions was calculated first in each family constellation (i.e., number of firstborns in two-child families, three-child families, number of second borns in two-child families, three-child families, and so forth up to 12-child families). In addition, tabulation was made of ordinal position regardless of family size (i.e., number of firstborns, second borns, etc.). Chi squares for expected and observed frequencies were calculated by the Greenwood-Yule method, and other comparisons were made with use of Mainland's tables. Results indicated that eldest children from small families and youngest children from large families, as well as only children, were overrepresented in this sample.

---

### A. INTRODUCTION

Relationships between order of birth and personality variables have intrigued researchers since the early part of the century. A principal hypothesis has been that the ordinal rank of a child in the family, in interaction with the number, sex, and order of siblings, is linked to the frequency and type of mental illness that an individual may develop.

Early studies by Pearson (11) and others showed increased incidence of insanity, criminality, and tuberculosis in unduly large families and among firstborns. Greenwood and Yule (4), however, criticized these studies for statistical bias, and presented a method for accurate estimation of mean

---

\* Received in the Editorial Office on May 23, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Requests for reprints and statistical data should be sent to Bernard F. Riess at the address shown at the end of this article. The authors wish to thank Mr. Michael Grossinger for his assistance in tabulating and analyzing the data.

family size. Applying their corrections to previous data, they found no relationship between insanity and either birth order or family size.

Since this earlier work, numerous researchers (2, 3, 7, 10, 14) have failed to discover any significant relationship between any aspect of birth order or family size and mental illness. This is particularly surprising because of the accumulation of social psychological evidence (12, 13) and clinical opinion (1, 6, 15) that different temperaments and predilections to psychopathology and characteristic of different birth positions.

There are several possible explanations for the lack of significant findings. Failure to find the predicted relationships may be due to the small size and atypical characteristics of samples; Schooler's data (14), for example, come from 63 hospitalized psychotic females, Patterson's (10) from 442 hospitalized schizophrenics, and Chakraborty's (2) are based upon data from 494 neurotic and psychotic outpatients in India, where the different birth positions may have quite different effects than in Western cultures.

The research of Ian Gregory (5) represents substantial methodological advance. His careful reanalysis of the Malzberg (7) and Norton (9) data, which the original authors thought demonstrated no differences in incidence rates among only children, eldest, youngest, or intermediates, showed that these negative findings were a function of the facts that the sexes had not been separated, although sex has been shown to interact with ordinal rank, and that data on eldest and only children were not separated. When the Greenwood-Yule corrections were made and the data reworked, the following significant findings appeared: for the schizophrenics and manic-depressives in Malzberg's sample there were more youngest children from sibships of four or more, and of the manic-depressives, eldest and youngest children were significantly overrepresented. In Norton's 2400 psychiatric outpatients, reanalysis revealed an excess of youngest and a deficiency of eldest and intermediate children.

The research reported here applies Gregory's methodology to a large, relatively homogeneous population of patients from a middle-income, outpatient, psychiatric clinic in New York City and is based on the largest such population yet studied in this context. The intention was to explore relationships between birth order and such other personality variables as sex, diagnosis, and sibling rank.

The purpose of this study was to demonstrate that birth order research should be done with a sufficiently large population so as to be able to factor out such variables as family size, sex, and other characteristics that

may come up in the course of explorative research. As a result of this type of statistical probing, hypotheses may be generated about the role of parental expectations, pressures, sibling rivalry, etc. in the etiology of patient problems.

## B. METHOD

Two thousand four hundred and seventy-four consecutive terminated cases, 1138 males and 1335 females, at the Postgraduate Center for Mental Health between 1947 and 1962 were collected, and the following data obtained: patient's sex; marital status; birth order in family; number, sex, and age of siblings; diagnosis on admission; patients' age and marital status. Twins, half siblings, and dead siblings where birth order information was not available from clinic records were omitted. From these data the frequency of each ordinal position within each sibship (for instance, first of two children was designated 1/2) and regardless of family size (i.e., number of firstborns, second borns, etc.) was tabulated for the total sample. Data were also organized by sex and by family size of patients. Families with one-three sibs were considered small, and those with four-12 considered larger. Since average patient age was computed to be 30.12 years, most families could be considered completed at the time of acceptance of the patient in the clinic.

Gregory's (5) formulation of the Greenwood-Yule reconstruction of family size and birth order was used to compute expected and observed frequencies within sibships, between sibships, and regardless of family size. Expected frequencies were obtained by summing the weighted frequencies from the bottom up, according to Gregory's method. Chi squares were calculated for each of the three categories above. Additional comparisons of observed frequencies were then made by means of Mainland's Fourfold Contingency Tables (8). Included were comparisons of the various birth positions both within and between small and large families, diagnosis compared with birth order, and sex of patient compared with birth order and family size.

## C. RESULTS AND DISCUSSION

Chi squares for ordinal position in relation to size of family (called "sibship" by Gregory), for ordinal position regardless of family size, and for position within each family size were calculated, and the results shown in Tables 1 and 2. Table 1, in which sibships were reconstructed according to the Greenwood-Yule methodology, shows that the distribution of sibship

TABLE 1  
RECONSTRUCTION OF FAMILY SIZE AND BIRTH ORDER: SIBSHIPS OF 2473 PATIENTS  
AT THE POSTGRADUATE CENTER FOR MENTAL HEALTH, NEW YORK, 1945-1962

Size of sibship	Frequency of subship from sample	Weighted frequency	Expected frequency of each birth order	Observed frequency of each birth order	Observed-expected frequency
1	353	353	1133.3	1137	.012
2	931	465.5	780.3	751	1.10
3	550	183.3	314.8	277	4.54
4	290	72.5	131.5	155	4.20
5	158	31.6	68.8	68	.009
6	97	16.1	27.2	48	15.91
7	48	6.8	11.3	19	5.25
8	25	3.1	4.5	8	2.72
9	6	.6	1.4	4	4.83
10	7	.7	.8	1	.05
11	2	.1	.1	0	.1
12	1	0	0	0	0

sizes in our patient population is significantly different from that to be expected in a "normal" distribution.<sup>2</sup> The contribution of each sibship size to the overall significance will be shown later. In Table 2 are shown the variations in ordinal positions within families of varying numbers. The chi squares show the significance of differences between observed and expected samples within each family size.<sup>3</sup> All differences were significant except for the two-sib family. In other words, regardless of family size, a family with a patient has a different frequency of ordinal positions from a nonpatient family. Thus, according to our data, children of different birth orders (firstborns, second borns, etc.) and positions within the family (1/2, 2/4, etc.) are not equally likely to become psychiatric patients, but some orders and ordinal positions are overrepresented in our population. As will be discussed below, this finding relates to the size of the family.

Data were then broken down in terms of whether the patient came from a small (one-three sibs) or large (four-12 sibs) family. The proportions of eldest and youngest children were compared via Mainland's tables both within and across family sizes. The results are presented in Table 3. The following comparisons were significant at the .01 level, with the larger frequency given first: eldest in small families compared with eldest in large families, eldest in small families compared with youngest in small

<sup>2</sup>  $\chi^2$  (obt.) = 38.72.  $\chi^2 = 23.21$ ;  $df = 10$ ;  $p < .01$ .

<sup>3</sup> Chi square values required for significance for sibship sizes 3 through 7, respectively, are as follows: 3.84, 9.2, 11.34, 13.38, 15.07.



TABLE 2  
FREQUENCY OF VARIOUS BIRTH POSITIONS WITHIN EACH FAMILY SIZE

Size of sibship	Ordinal position in family	Observed frequency	Expected frequency	Chi square observed	Significance level
2	1	480	464.5	.9	N.S.
	2	451	"		
3	1	203	183.3	4.15	.05
	2	183	"		
		164	"		
4	1	52	72.5	11.6	.01
	2	73	"		
	3	72	"		
	4	93	"		
5	1	30	31.8	14.22	.01
	2	22	"		
	3	21	"		
	4	41	"		
	5	44	"		
6	1	13	16.2	16.63	.01
	2	13	"		
	3	12	"		
	4	15	"		
	5	13	"		
	6	31	"		
7	1	3	6.8	24.35	.01
	2	5	"		
	3	5	"		
	4	1	"		
	5	8	"		
	6	9	"		
	7	17	"		

families, youngest in large families compared with eldest in large families (that is, there were more eldest children from small families than from large families). Also within small families, more eldest children were patients, while within large families more youngest children came for treatment. No significant differences was found between frequencies of youngest children in

TABLE 3  
FREQUENCIES OF ELDEST AND YOUNGEST CHILDREN IN SMALL AND LARGE FAMILIES

Family size	Ordinal position					
	Eldest			Youngest		
	Male	Female	Total	Male	Female	Total
Small ( $N = 1834$ )						
One-three siblings	469	566	1035	265	350	615
Large ( $N = 634$ )						
Four-12 siblings	46	61	107	90	103	193

small families and youngest children in large families, a finding which deserves investigation along with the equal incidence rate in two-sib families. In sum, these data indicate that a large percentage of patients at the Postgraduate Center were either eldest children from small families or youngest children from large families, a finding which holds both within each subgroup and between them.

This finding is the most significant in the data. That earlier studies did not find this relationship may be due to their small or atypical samples.

In accordance with the exploratory nature of this study, hypotheses about other birth order related phenomena in the data were generated and tested. Among these were relationships between sex, ordinal position, and family size; characteristics of only children and firstborns; and sociological and diagnostic comparisons. Mainland's tables (8) were used to establish significance. Analysis indicated a preponderance of females in the first three birth positions regardless of family size and a higher frequency of females in small families (1010 females *vs.* 823 males, significant at .01). The overall ratio of females to males in the sample is 1335:1138, a moderately significantly (.05) larger number of females.

Several other sociological findings emerged from the data. When divorce rates for firstborns and later borns were compared, no significant difference was found (151 divorced firstborns *vs.* 142 later borns) in contradiction to Schachter's (13) finding that more firstborns terminated their marriages. Divorce of parents does seem to affect patient frequencies, however; significantly more Postgraduate Center patients came from broken homes (1316 *vs.* 1157 from intact homes, significant at .01). Another finding regarding marital status is the preponderance of never-married persons in the patient population (1373 *vs.* 750 married, significant at .01). This finding, based on data from 1945-1962, contradicts the report in the August 1970 HEW publication entitled "Selected Symptoms of Psychological Distress" (16). Data from the National Health Survey carried out in 1970 discussed therein found never-married persons to have lower symptom rates than any other group, with 32 of the 48 symptoms in their checklist lower than expected.

Numerous studies in the past have singled out only children as prime candidates for neurosis. For this sample, there were significantly more only children than children in any other birth position: 15% *vs.* 88% multiples. As Gregory and others have contended, only children do seem to be different from firstborn children with siblings and for this reason should be analyzed separately in birth order investigations. The only significant

finding relating birth order to initial diagnosis in this sample was a comparison between firstborns and only children in which only children comprised more of the patients diagnosed passive-aggressive, psychoneurotic, homosexual, depressed, and obsessive-compulsive, the most common diagnoses at the Postgraduate Center during those years.

As part of our preliminary investigations into reasons for the distributions presented in this paper, an hypothesis of König (6) was tested. He claims that the personalities of first-, second-, and third-born children have discrete characteristics, and that the higher birth orders repeat these character types, fourth-borns being similar to firstborns, fifths to seconds, and sixths to thirds. By use of initial diagnoses as a basis for differentiating some aspects of personality, no significant differences were found between the frequencies of first and fourth, second and fifth, and third and sixth borns in any diagnostic category. When diagnoses were grouped into the larger categories of psychotic, neurotic, character disorder, and homosexuality, significant differences did emerge between first, second, and third born children as König would predict. These findings are shown in Table 4.

TABLE 4  
DIFFERENCES IN DIAGNOSES OF FIRST-, SECOND-, AND THIRD BORN CHILDREN

Diagnosis	Firstborn N = 1137		Second born N = 751		Third born N = 277	
	N	%	N	%	N	%
Psychotic	54	5	71	9 <sup>a</sup>	57	20
Neurotic	82	7	119	16	86	31
Character disorder	89	8	134	18	98	35
Homosexuality	5	.04	8	1 <sup>a</sup>	11	4

<sup>a</sup> All comparisons but these are significant at the .01 level.

The next step after establishing that order of birth in relation to family size leads children of certain birth positions to become mental patients more frequently than others must now be investigated in depth. Case study is certainly indicated. As suggested by our preliminary work with König's hypothesis, the future directions of this research should be in causation and eventual prevention of whatever traumas and conflicts are associated with certain places in the family constellation.

#### REFERENCES

1. ANSBACHER, H. C., & ANSBACHER, R. R. *The individual Psychology of Alfred Adler*. New York: Basic Books, 1956.

2. CHAKRABORTY, A. Birth order and mental illness. *Brit. J. Soc. Psychiat.*, 1969-70, 3(4), 231-236.
3. DAMRIN, D. E. Family size and sibling age, sex, and position as related to certain aspects of adjustment. *J. Soc. Psychol.*, 1949, 29, 93-102.
4. GREENWOOD, M., & YULE, G. U. On the determination of size of family and of the distribution of character in order of birth from samples taken through members of sibship. *J. Royal Stat. Soc.*, 1914, 77, 197-199.
5. GREGORY, I. An analysis of familial data on psychiatric patients: Parental age, family size, birth order, and ordinal position. *Brit. J. Prev. Soc. Med.*, 1958, 12, 42-59.
6. KONIG, K. Brothers and Sisters. New York: Anthroposophic Press, 1963.
7. MALZBERG, B. Social and Biological Aspects of Mental Disease. Utica, N. Y.: State Hospitals Press, 1940.
8. MAINLAND, D., HERRERA, L., & SUTCLIFFE, M. Tables for use with binomial samples. Department of Medical Statistics, New York University College of Medicine, New York, New York, 1956.
9. NORTON, A. Unpublished data, State Hospitals Press, Utica, New York, 1957.
10. PATTERSON, R., & ZEIGER, T. Ordinal position and schizophrenia. *Amer. J. Psychiat.*, 1941, 98, 455-458.
11. PEARSON, K. On the handicap of firstborns. Volume No. 10, Eugenics Laboratory Section Service, London, England, 1914.
12. SAMPSON, E. The study of ordinal position: Antecedents and outcomes. In B. A. Maher (Ed.), *Progress in Experimental Personal Research* (Vol. 2). New York: Academic Press, 1965.
13. SCHACTER, S. The Psychology of Affiliation. Stanford, Calif.: Stanford Univ. Press, 1959.
14. SCHOOLER, C. Birth order and schizophrenia. *A.M.A. Arch. Gen. Psychiat.*, 1961, 4, 91-97.
15. TOMAN, W. Family Constellation. New York: Springer, 1961.
16. U. S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE. Selected symptoms of psychological distress. Series 11, Number 37, National Center for Health Statistics, Rockville, Maryland, 1970.

*Postgraduate Center for Mental Health*  
*124 East 28 Street*  
*New York, New York 10016*

## IMPLICIT TIME CONSTRAINTS IN THE MEASUREMENT OF PRODUCTIVE THINKING\*

*Northwestern University and Pennsylvania State University*

---

ROBERT J. MENGES AND BELINDA JOHNSON

---

### SUMMARY

A test of productive thinking was taken by 265 university freshmen by mail during the summer prior to their matriculation and by 89 other freshmen, members of a subject pool, during the academic year. Although no time limits were made explicit in either case, the subject pool group reported significantly less time spent on the questionnaire. Apparently implicit time constraints differed for the two groups. Nevertheless, the relationship between ability and productive thinking was not different in the two groups.

---

### A. INTRODUCTION

Researchers have been concerned with the effects of the setting in which data on productive thinking are gathered. Several investigators have contrasted test-like administration with game-like administration. Dentler and Mackler (1) and Vernon (6) found more high scores under the latter (relaxed) conditions, whereas Kogan and Morgan (2) found no differences.

Time constraints may be intentionally manipulated by the investigator and made explicit for subjects. But they may also be implicit in the data collection setting, existing in the minds of subjects independent of the experimenter's intentions. The latter possibility is explored here.

The present study provides comparison of mail respondents with subject pool respondents, all of whom were ostensibly under "relaxed" conditions.

### B. PROCEDURE

Data were gathered as part of a study of freshmen at a large midwestern state university (3). Productive thinking was assessed by questions used by Wallach and Wing (9). The instrument has four three-part items. The first item, "uses," lists three objects and for each object asks for "all the different ways you can think of in which the object might be used." The

---

\* Received in the Editorial Office on May 25, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.



second, "pattern meaning," reproduces three patterns and for each pattern asks for "all the different things you can think of that each pattern might suggest." The third, "similarities," names three pairs of objects and for each pair asks for "all the different ways you can think of in which the two objects might be alike." The fourth, "line meanings," reproduces three line sketches and for each sketch asks for "all the different things you can think of that each complete line might suggest." To summarize: there were presented in mixed order two verbal tasks (uses and similarities) and two visual tasks (pattern meanings and line meanings).

The sample included every seventh new student admitted for Fall, 1970. Each was sent a questionnaire during the summer preceeding matriculation and offered \$3.00 for completing it. Acceptable questionnaires were returned by 265 students (40%). During October, 100 members of the introductory psychology course subject pool were scheduled for a two-hour session where they completed the same questionnaire. All were new students, and none had been mailed a questionnaire. For attending that session they received two hours of credit toward their five-hour subject pool requirement. Eighty-nine students provided scorable responses.

Neither group completed the questionnaire under an explicit time limit. Instructions to the mail group made no mention of time. Oral instructions to the subject pool group included the following: "Instructions are in the booklet. Please take as much time as you wish to take. As you leave, we will give you a mimeographed description of the purpose of the research and will be glad to answer other questions." Furthermore, no decisions about any individual's future depended on his performance. Thus, the contrast is between responses on a questionnaire filled out at home for a small cash payment and responses on a questionnaire completed with a group in a classroom to satisfy part of a requirement to participate in research. Both settings were "relaxed," and neither imposed explicit time constraints. Therefore, no difference in performance was expected.

### C. RESULTS AND DISCUSSION

A highly significant difference ( $p < .002$ ) on time spent on the questionnaire was found, as shown in Table 1, with the mail respondents taking about twice as long. As might be expected, the mail group had higher productive thinking scores ( $p < .02$ ). Also, the mail group checked more nonacademic activities ( $p < .05$ ). On other measured characteristics the groups did not differ: American College Testing Program Composite Score (ACT-C) taken during the college admission process, School and College Ability Test Verbal

TABLE 1  
MEAN SCORES FOR MAIL AND SUBJECT POOL RESPONDENTS

Variable	M	F	Mail Total	t	p	M	Subject Pool F	Subject Pool Total	t	Mail vs. subject pool p	p
ACT-C											
N	139	126	265			33	36	89			
M	27.06	25.52	26.33	4.20	.002	27.51	25.08	26.33	4.29	.002	.53
SD	2.68	3.28	3.07			2.35	2.97	2.86			ns
SCAT-V											
N	138	126	264			50	35	85			
M	31.43	32.24	31.82	.70	ns	31.66	32.89	32.16	.71	ns	.31
SD	9.23	9.43	9.32			8.63	6.62	7.84			ns
HSPR											
N	139	126	265			53	36	89			
M	84.85	87.52	86.12	1.79	.10	85.37	84.89	85.29	.25	na	.55
SD	13.45	10.50	12.19			11.62	14.10	12.60			ns
GPA											
N	138	124	262			53	36	89			
M	3.87	4.06	3.96	2.14	.05	3.67	4.18	3.88	3.23	.002	.90
SD	.72	.71	.72			.73	.76	.78			ns
Time											
N	139	124	263			53	36	89			
M	1.10	1.13	1.12	.40	ns	.68	.68	.68	.07	ns	6.42
SD	.60	.68	.64			.17	.16	.16			.002
Prod thkg											
N	139	126	265			53	36	89			
M	63.06	78.12	70.22	4.22	.002	61.15	63.44	62.08	.69	ns	2.46
SD	25.27	32.71	29.95			15.33	15.45	15.33			.02
Non-ac											
N	139	126	265			53	36	89			
M	11.14	12.98	12.01	2.76	.01	9.58	12.28	10.67	2.65	.02	2.05
SD	5.32	5.52	5.48			4.28	5.29	4.87			.05

Note: ACT-C = American College Testing Program Composite Score; SCAT-V = School and College Ability Test Verbal Score; HSPR = high school percentile rank; GPA = first semester college freshman grade point average; Time = minutes for completing questionnaire; Prod thkg = score on productive thinking test; Non-ac = number of nonacademic activities.

Score (SCAT-V) taken during freshman orientation week, high school percentile rank (HSPR), and first semester freshman college grade point average (GPA). Both groups were representative of all freshmen on these academic measures.

Although there is no way of knowing if estimates of time spent on the questionnaire were accurately reported by the mail respondents, it was possible for the investigators to verify reports of the subject pool respondents. To our surprise we found systematic overestimation, probably in order to ensure getting credit for two hours of experimental participation. Thus the difference between groups in time, though not in scores, would be less if it were based on student's self-reports rather than on the researcher's notations.

These results suggest that there were important differences between conditions. Apparently there were constraints implicit in the subject pool classroom to finish quickly, even at the cost of performance quality.

A second question is of interest, once we grant that implicit time constraints depressed the level of responding for the subject pool group: Does that depression also affect relationships between variables? For example, is the correlation between productive thinking and ability higher in the subject pool group than in the mail group because of the implicit time constraints?

Wallach and Kogan (8) argued that the independence of intelligence from creativity is maximized when the latter is tested under relaxed conditions. Van Mondfrans *et al.* (5) found lower correlations with *IQ* when creativity was assessed under relaxed conditions. However Vernon (6), Kogan and Morgan (2), and Sherwood (4) found essentially zero correlations between intelligence (ability) and productive thinking measures *regardless* of administration conditions. After reviewing a number of such studies, Wallach (7) concludes "a permissive context for assessing ideational fluency is not necessary to demonstrate its independence from intelligence" (p. 14).

The present data support that conclusion. Table 2 presents intercorrelation matrices separately for males and females. Tests were made to determine if members of each pair of coefficients (mail *versus* subject pool) were different. Differences were found for only two pairs, both involving males. Correlations were significantly higher for mail than for subject pool males (a) between HSPR and ACT-C ( $p < .05$ ) and (b) between HSPR and Scat-V ( $p < .02$ ).

There is no evidence that productive thinking is related to other variables differently in the mail group than in the subject pool group. However, HSPR does function differently. For males it is more highly correlated with

TABLE 2  
INTERCORRELATIONS OF SELECTED VARIABLES FOR MALE AND FEMALE MAIL<sup>a</sup> AND  
SUBJECT POOL<sup>b</sup> RESPONDENTS

Variable	ACT-C	SCAT-V	HSPR	GPA	Prod thkg	Non-ac
<i>Males</i>						
ACT-C						
Mail	—					
Subject pool						
SCAT-V						
Mail	.65	—				
Subject pool	.66					
HSPR						
Mail	.53*	.43**	—			
Subject pool	.25*	.04**				
GPA						
Mail	.34	.28	.31	—		
Subject pool	.18	.16	.26			
Prod thkg						
Mail	.10	.19	.06	.24	—	
Subject pool	.17	.08	.12	.18		
Non-ac						
Mail	.07	-.01	.11	.02	.07	—
Subject pool	.00	-.04	-.03	.25	.30	
<i>Females</i>						
ACT-C						
Mail	—					
Subject pool						
SCAT-V						
Mail	.69	—				
Subject pool	.58					
HSPR						
Mail	.47	.27	—			
Subject pool	.63	.32				
GPA						
Mail	.20	.12	.33	—		
Subject pool	.39	.35	.47			
Prod thkg						
Mail	.03	.06	.08	.07	—	
Subject pool	.20	.23	.04	.43		
Non-ac						
Mail	.21	.17	.34	.17	.27	—
Subject pool	-.03	.10	.03	-.04	-.04	

Note: ACT-C = American College Testing Program Composite Score; SCAT-V = School and College Ability Test Verbal Score; HSPR = high school percentile rank; GPA = first semester college freshman grade point average; Time = minutes for completing questionnaire; Prod thkg = score on productive thinking test; Non-ac = number of nonacademic activities.

<sup>a</sup> N for each variable ranged from 137 to 139 for males and 124 to 126 for females.

<sup>b</sup> N for each variable ranged from 50 to 53 for males and 35 to 36 for females.

\* Members of this pair are significantly different ( $p < .05$ ).

\*\* Members of this pair are significantly different ( $p < .02$ ).

both ability measures in the mail group than in the subject pool group. This finding is difficult to interpret, since the groups did not differ on HSPR. We might speculate that there is a component of "school success" for males wherein ability and achievement are closely related. If this component also includes a behavioral pattern of compliance, we might expect it to be more prominent among those who return questionnaires (the higher correlations) than among those who attend appointments for a required subject pool (the lower correlations). A more direct test of this speculation would be provided by use of personality instruments that include compliance or conformity scales.

#### REFERENCES

1. DENTLER, R. A., & MACKLER, B. Originality: Some social and personal determinants. *Behav. Sci.*, 1964, 9, 1-7.
2. KOGAN, N., & MORGAN, F. T. Task and motivational influences on the assessment of creative and intellectual ability in children. *Genet. Psychol. Monog.*, 1969, 80, 91-127.
3. MENGES, R. J. Academic ability, nonacademic accomplishments, and ideational productivity in high-risk and regularly admitted college students. Final Report, Project No. O-E-121, Grant No. OEG-5-70-0040 (509), United States Office of Education, Washington, D.C., August, 1972.
4. SHEERWOOD, D. W. The differential effects of assessment context and scoring method on creativity performance in children. Doctoral dissertation, Duke University, Durham, North Carolina, 1968.
5. VAN MONDEFRAANS, A. P., FELDHEUSEN, J. F., TREFFINGER, D. J., & FERRIS, D. R. The effects of instructions and response time on divergent thinking test scores. *Psychol. Sch.*, 1971, 8, 65-71.
6. VERNON, P. E. Effects of administration and scoring on divergent thinking tests. *Brit. J. Educ. Psychol.*, 1971, 41, 245-257.
7. WALLACE, M. A. *The Intelligence/Creativity Distinction*. New York: Gen. Learn. Press, 1971.
8. WALLACE, M. A., & KOGAN, N. *Modes of Thinking in Young Children: A Study of the Creativity-Intelligence Distinction*. New York: Holt, Rinehart & Winston, 1965.
9. WALLACE, M. A., & WING, C. W., JR. *The Talented Student: A Validation of the Creativity-Intelligence Distinction*. New York: Holt, Rinehart & Winston, 1969.

*The Center for the Teaching Professions*  
*Northwestern University*  
*Evanston, Illinois 60201*



## BIRTH ORDER, PERSONALITY DEVELOPMENT, AND VOCATIONAL CHOICE OF BECOMING A CARMELITE NUN\*

*Rhode Island College*

---

PHILIP S. VERY,<sup>1</sup> ROBERT B. GOLDBLATT, AND VINCENT MONACELLI

---

### SUMMARY

This study was undertaken to determine whether there is a relation between birth order, as described by Konig and Very, and becoming a Discalced Carmelite nun. Questionnaires were sent to the United States population of fully professed Discalced Carmelite nuns, and return data compared with the United States Vital Statistics for female live births by birth order.

Chi square analysis showed a significant difference ( $p = .01$ ) supporting the hypothesis that the distribution of Carmelite nuns by birth order is different from the expected distribution. A greater proportion of third borns than expected became nuns. Other birth orders came closer to chance. It appears that there is a relationship between birth order and personality development, and that personality is a factor in an individual's vocational choice.

---

### A. INTRODUCTION

This study is to demonstrate that the order of birth of an individual affects his personality, which in turn has an effect upon his vocational choice. Specifically, this study concerns itself with the birth order of individuals choosing to become Carmelite nuns.

Ever since Frank Parsons' work *Choosing a Vocation* (6) appeared in the early part of this century, individuals have been concerned about the process of vocational choice. According to Parsons (p. 3):

The wise selection of the business, profession, trade, or occupation to which one's life is to be devoted and the development of full efficiency in the chosen field are matters of deepest moment to young men and the public. These vital problems should be solved in a careful scientific way.

---

\* Received in the Editorial Office on June 1, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Reprints available from the first author at the address shown at the end of this article.

Since Parsons' work such men as Eli Ginzberg (7) have formulated vocational concepts and theories to explain the process of selecting a vocation. Recently, more emphasis has been placed on the role of personality in occupational choice (1, 2, 3, 5, 8).

This study assumes that birth order is an important factor in personality development and that personality development in turn affects vocational choice. This position does not hypothesize that birth order *per se* affects vocational choice, but rather that the development of the personality in the young child is affected by the individual birth order, and this resultant personality structure serves as a factor in a person's moving toward a vocation that will gratify his particular personality needs. The birth order theories of Konig (4) and Very (9, 10, 11) were utilized as the criteria for formulating the hypothesis. Konig and Very postulate four basic personality types: (a) the "only child", (b) the "firstborn" who is the first of two or more siblings, (c) the "second born," and (d) the "third born." According to these theories if the family exceeds three children in number, the basic triadic pattern of first, second, and third born personality types is repeated in the fourth, fifth, and sixth child, respectively, and so forth.

The firstborn of two or more children, according to Konig and Very, tends to have a high sense of responsibility, to be duty bound, and to have more drive and perserverance than the other siblings. He can generally be called a "doer."

The second born tends to be less duty bound and could generally be characterized as easy-going, cheerful, stubborn, light-hearted, casual, and unconcerned.

The third born is the most sensitive, withdrawn, emotional, and distrustful of the three. He tends to respond more from "inner" urges than from the outer demands.

The only child in America tends to be a combination of the first and third, being less dynamic and aggressive than the first, and less withdrawn and sensitive than the third.

Very (9, 10, 11) has demonstrated that birth order is a significant factor in affecting vocational choice. His predictions of differing personality traits of first and second borns, predisposing them towards careers of law and becoming a beautician, respectively, have been upheld.

This study attempts to analyze the inherent personality and attitudinal characteristics of those individuals who become Carmelite nuns. It is hypothesized that individuals drawn to this vocation would tend to be those who are somewhat more withdrawn and less social than others. The pattern of

personality which would seem most appropriate for this vocation closely parallels the pattern of personality described by Konig and Very as the third born. Specifically, then, it is predicted that the greater proportion of third borns than expected by chance would choose this vocation.

## B. METHOD

### 1. *Subjects*

Sixty-three monasteries were selected from the Official Catholic Directory which listed a total population of approximately 700 fully professed nuns. The total sample consisted of 583 Discalced Carmelite nuns from 56 monasteries in the United States, constituting approximately 85 percent of the total Carmelite nun population.

### 2. *Questionnaire*

The questionnaire used required Ss to indicate their own age, sibling ranking, and sex plus the birth years of their brothers and sisters. Only siblings born alive were to be reported.

### 3. *Procedure*

The data were obtained by sending a letter to the prioress in each of 63 monasteries, together with the appropriate number of questionnaires and a return-addressed stamped envelope. Returned questionnaires were then sorted into four classifications: only, first, second, and third born.

## C. RESULTS

Because the United States Vital Statistics for female live births by birth order were not available for births prior to 1917, only the segment of the nun population ( $N = 380$ ) that included live births from 1917 to 1943 was used in this study, so that a true comparison could be made.

In order to prevent confounding of the triadic pattern variable with the birth order variable itself, individual chi square analyses were first done on each of the triads. Percentages of observed and expected frequencies for actual first, second, and third borns were compared, then fourth, fifth, and sixth, then seventh, eighth, and ninth. Analysis was not continued beyond the ninth position, since the numbers involved after that position were too small for statistical comparison. The distribution of Carmelite nuns for all three triads was statistically significantly different from the expected distribution at the one percent level of significance. In all three triads, third,

sixth, and ninth borns were more represented in the distribution than would be expected by chance for that position.

According to the theories described previously, there is a triadic pattern of first, second, and third born positions which repeats itself through succeeding siblings. Therefore, fourth borns were classified as firstborns, fifth borns as second borns, etc. Also, the firstborn of a family of two or more was classified as a firstborn, whereas the firstborn of a family that had no further children was classified as an only child. On the basis of the above analyses, the triads were combined into four categories of "only," "first" (first, fourth, seventh), "second" (second, fifth, eighth), and "third" (third, sixth, ninth).

Of the 380 Ss, 15 were "only," 149 were "first," 106 were "second," and 110 were "third," yielding 3.9%, 39.2%, 27.9%, and 28.9%, respectively.

The United States Vital Statistics for female live births by birth order for the years 1917-1943 were as follows: 9.6% only, 36.9% firstborns, 32.4% second borns, and 21.1% third borns. A chi square analysis of the percentages of the observed and expected frequencies was performed ( $\chi^2 = 25.98$ ). The distribution of Carmelite nuns by birth order was statistically significantly different from the expected distribution of the general population by birth order, which was confirmed at the one percent level of significance with three degrees of freedom. The reason for this significance is illustrated in Figure 1, which shows that the prediction concerning the third borns in the nun population was upheld: i.e., there were more third borns (28.9%) than expected (21.1%); there were slightly more firstborns (39.2%) in the nun population than expected (36.9%); there were fewer second born nuns (27.9%) than expected (32.4%); and the percentage of only children in the nun population (3.9%) was less than half the expected frequency (9.6%).

#### D. DISCUSSION

The results obtained tend to support the theory that there is a relationship between birth order and personality traits affecting vocational choice. The personality characteristics of Carmelite nuns appear to be most similar to those of third born individuals, since a significantly greater proportion of nuns than expected are in fact third born. However, certain variables might confound and reduce the quantity of significance. While one purpose of this study is to indicate that people tend to gravitate to professions that match their personality and attitudinal predispositions, in actual fact this is not always the case. For many reasons people may enter and remain in a profession that is not personally appealing to them. Such people would, of course,

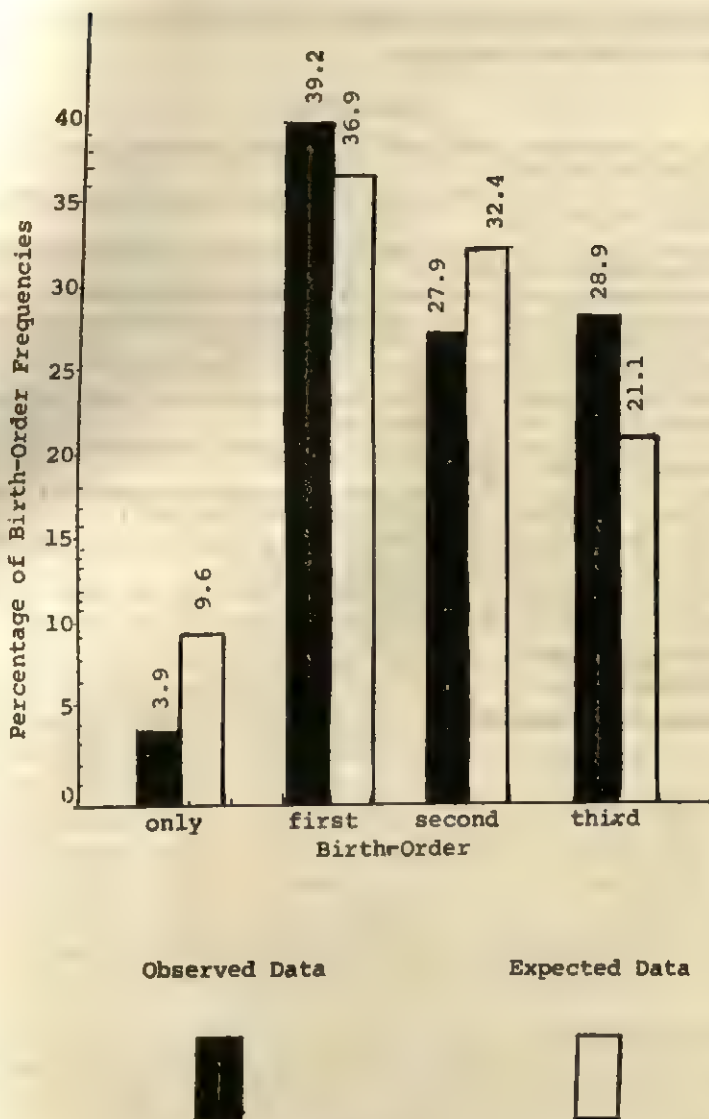


FIGURE 1  
PERCENTAGES OF OBSERVED DATA (NUNS) AND EXPECTED DATA  
(U. S. VITAL STATISTICS) BY BIRTH DATA

Firstborn includes 4th, 7th, 10th, and 13th; second born includes 5th, 8th, 11th, and 14th; third born includes 6th, 9th, 12th, and 15th.



weaken efforts to analyze relationships between birth order, personality characteristics, and vocational choice.

## REFERENCES

1. BOHN, M. Psychological needs related to vocational personality types. *J. Counsel. Psychol.*, 1966, **13**, 306-309.
2. CHILDERS, R., & WHITE, W. The personality of select theological students. *Person. & Guid. J.*, 1966, **44**, 507-510.
3. GALINSKY, M., & FAST, I. Vocational choice as a focus of the identity search. *J. Counsel. Psychol.*, 1966, **13**, 89-92.
4. KONIG, K. Brothers and Sisters. Blauvelt, N. Y.: St. George Books, 1963.
5. OPPENHEIMER, E. A. The relationship between certain self-constructs and occupational preferences. *J. Counsel. Psychol.*, 1966, **13**, 191-197.
6. PARSONS, F. Choosing a Vocation. Boston: Houghton Mifflin, 1909.
7. PETERS, H., & HANSEN, J., Eds. Vocational Guidance and Career Development: Selected Readings. New York: Macmillan, 1966.
8. ROE, A., & SIEGELMAN, M. The Origins of Interests. Washington, D. C.: Amer. Person. & Guid. Assoc., 1964.
9. VERY, P. S., & PRULL, R. W. Birth order, personality development, and the choice of law as a profession. *J. Genet. Psychol.*, 1970, **116**, 219-221.
10. VERY, P. S., & VAN HINE, N. P. Effects of birth order upon personality development of twins. *J. Genet. Psychol.*, 1969, **114**, 93-95.
11. VERY, P. S., & ZANNINI, J. A. Relation between birth order and being a beautician. *J. Appl. Psychol.*, 1969, **53**(2), 149-151.

*Psychology Department*  
*Rhode Island College*  
*Providence, Rhode Island 02908*

## DEVELOPMENT OF A PSYCHOLOGICAL EFFECTIVENESS SCALE\*<sup>1</sup>

*Iowa State University*

---

CHARLES A. POE<sup>2</sup>

---

### SUMMARY

The present study provided evidence to support the hypothesized trait of psychological effectiveness. The findings indicated that psychological effectiveness can be assessed psychometrically. The Psychological Effectiveness Scale which was developed was observed to have high internal consistency reliability, and to have discriminated among concepts of personality which represented distinctive levels of effective behavior. The results of the present study indicate that the scale has sufficient preliminary reliability and validity evidence to warrant further study and development.

---

### A. INTRODUCTION

In recent years there has been growing interest in the study of normal and optimal personality (e.g., 4, 5, 6, 7, 9, 20, 21, 22, 24, 25). Though the intensity of interest has been increasing, the focus of the study has been somewhat diffuse as reflected by the numerous concepts utilized: e.g., competence (23), integration (16), positive mental health (8), psychological effectiveness (10), and optimally adjusted person (19). Similarly, a number of variables thought to characterize the normal or optimal personality have been hypothesized: e.g., capacity for interpersonal relationships (17), effective organization of work (3), extension of self (1), inner-directedness (18), and openness to experiences and realistic perceptions (15). The present author, in accord with these previous authors, hypothesized a trait, or traits, of psychological effectiveness. It was further hypothesized that this trait, or

---

\* Received in the Editorial Office on June 4, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This report is based on the following paper: Poe, C. A., & Matias, R. A comparison of conceptions of normal, optimal, and abnormal personality. Paper presented at the meeting of the American Psychological Association, Washington, D. C., September, 1969. Financial support for the study was provided by the Student Counseling Service, Iowa State University.

<sup>2</sup> The author is now at the Psychology Service, Veterans Administration Center, Hampton, Virginia 23667. Copies of the Psychological Effectiveness Scale may be obtained from the author at this address. Appreciation is expressed to John Menne for statistical consultation.

traits, could be assessed psychometrically: i.e., would meet minimal preliminary criteria of reliability and validity.

### B. METHOD

The technique selected for the measurement of psychological effectiveness was the semantic differential (13). However, to obtain specific item content judged more relevant to the measurement of psychological effectiveness than the original semantic differential items, trait names and behavioral phrases were selected from the literature on normal and optimal personality; the sources of the items are illustrated by the selected references at the end of this article. A second source of items was the responses by psychologists to incomplete sentence stems (e.g., a mentally healthy person feels . . . , a person who is normal believes . . . , etc.). From the trait names and behavioral phrases selected from the literature and responses of psychologists, the author constructed 50 semantic differential items. This method of scale construction meets minimal criteria for content validity (2, 12).

Utilizing the 50-item scale, 93 college students were instructed to assign meanings to selected personality concepts on the nine-point semantic differential items. Four concepts frequently encountered in the literature—abnormal, typical, mentally healthy, and ideal person—were selected because of their representation of different levels of effective functioning. Responses to the 50 items were summed to a single total score for each concept per subject, and the differences of the total scale scores for the personality concepts were compared by *t* tests.

To determine the internal consistency reliability of the items, coefficient alpha was computed for ratings of each of the four concepts.

To specify in a limited manner the content of the Psychological Effectiveness Scale, five items which correlated most highly with the total score ratings for the four personality concepts were selected for tabular presentation.

### C. RESULTS

The means of the 50-item Psychological Effectiveness Scale were computed for the four concepts and are presented in Table 1, along with standard deviations. All four concept means were significantly different from each other ( $p < .005$ ). Even by inspection, there appeared to be large differences between the mean ratings of the concepts.

The internal consistency reliability of the Psychological Effectiveness Scale was indicated by coefficient alpha which across the four concepts—

TABLE 1  
PERSONALITY CONCEPT DIFFERENCES IN PSYCHOLOGICAL EFFECTIVENESS RATINGS

Statistic	Abnormal	Personality concept		Ideal
		Typical	Mentally healthy	
Mean <sup>a</sup>	160.73	295.13	353.95	395.32
SD	41.43	48.27	48.06	37.77

<sup>a</sup> All means differ significantly with  $p < .005$ .

abnormal, typical, mentally healthy, and ideal person—was .92, .97, .95, and .95, respectively.

The five most discriminatory items across the four personality concepts are presented in Figure 1.

Item #	Mean Concept Rating <sup>a</sup>									
39. profits from	I	M	T			A				repeats
experience	:	:	:	:	:	:	:	:	:	mistakes
31. regrets his				A		T	M	I		makes his deci-
decisions	:	:	:	:	:	:	:	:	:	sions "work out"
32. denies				A		T		M	I	profits from
mistakes	1	2	3	4	5	6	7	8	9	mistakes
18. direction-				A		T		M	I	purposive
less	:	:	:	:	:	:	:	:	:	
20. makes good	I	M	T					A		makes poor
choices	:	:	:	:	:	:	:	:	:	choices

<sup>a</sup>"A" indicates mean item response for abnormal person, "T" for typical, "M" for mentally healthy, and "I" for ideal.

FIGURE 1  
ITEMS MOST DISCRIMINATORY BETWEEN PERSONALITY CONCEPTS

## D. DISCUSSION

The 50-item Psychological Effectiveness Scale was observed to have high internal consistency reliability in the rating of the personality concepts—abnormal person, typical person, mentally healthy person, and ideal person. Coefficient alpha was equally high when each of the four concepts was rated. Subsequently, a shorter 40-item form of the Psychological Effectiveness Scale was observed to be highly reliable also; Nowacki and Poe (11) reported that the ratings of concepts of mentally healthy males and females were high (alpha equaled .91). More recently, Poe (14) found alpha to be .91 in self-ratings.

As stated earlier, the method of constructing the Psychological Effectiveness Scale met minimal criteria for content validity (2, 12). Further support for the validity of the Psychological Effectiveness Scale was found in the differential ratings on the Scale, as expected, of the personality concepts, which represented distinctive levels of effective behavior—abnormal, typical, mentally healthy, and ideal personality.

In conclusion, the present study provided preliminary evidence to support the hypothesized trait of psychological effectiveness. The findings indicated that psychological effectiveness can be assessed psychometrically. These preliminary results suggest that the scale has sufficient reliability and validity to warrant further study and development.

## REFERENCES

1. ALLPORT, G. W. *Pattern and Growth in Personality*. New York: Holt, Rinehart & Winston, 1963.
2. AMERICAN PSYCHOLOGICAL ASSOCIATION. *Standards for Educational and Psychological Tests and Manuals*. Washington, D. C.: A. P. A., 1966.
3. BARRON, F. *Creativity and Psychological Health*. Princeton, N. J.: Van Nostrand, 1963.
4. FOREMAN, M. E. Some empirical correlates of psychological health. *J. Counsel. Psychol.*, 1966, **13**, 3-11.
5. HEARN, C. B., & SEEMAN, J. Personality integration and perception of interpersonal relationships. *J. Personal. & Soc. Psychol.*, 1971, **18**, 138-143.
6. HEISLER, V. Toward a process model of psychological health. *J. Counsel. Psychol.*, 1964, **11**, 59-62.
7. HOGARTY, G. E., & KATZ, M. M. Norms of adjustment and social behavior. *Arch. Gen. Psychiat.*, 1971, **25**, 470-480.
8. JAHODA, M. *Current Concepts of Positive Mental Health*. New York: Basic Books, 1958.
9. KARON, B. P., & O'GRADY, P. Quantified judgments of mental health from the Rorschach, TAT, and clinical status interview by means of a scaling technique. *J. Consult. & Clin. Psychol.*, 1970, **34**, 229-235.
10. LEVINE, L. S., & KANTOR, R. E. Psychological effectiveness and imposed social position. *Person. & Guid. J.*, 1962, **40**, 418-425.



11. NOWACKI, C. M., & POE, C. A. The concept of mental health as related to sex of person perceived. *J. Consult. & Clin. Psychol.*, 1973, **40**, 160.
12. NUNNALLY, J. C. *Psychometric Theory*. New York: McGraw-Hill, 1967.
13. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The Measurement of Meaning*. Urbana, Ill.: Univ. Illinois Press, 1957.
14. POE, C. A. Internal consistency of a psychological effectiveness scale. *Psychol. Rep.*, in press.
15. ROGERS, C. R. The concept of the fully functioning person. *Psychother.: Theory, Res. & Prac.*, 1963, **1**, 17-26.
16. SEEMAN, J. Toward a concept of personality integration. *Amer. Psychol.*, 1959, **14**, 633-637.
17. SHOEN, E. J. Toward a concept of the normal personality. *Amer. Psychol.*, 1957, **12**, 183-190.
18. SHOSTROM, E. L. *Personal Orientation Inventory Manual*. San Diego, Calif.: Educ. & Indus. Test. Serv., 1966.
19. SIEGELMAN, E., BLOCK, J., BLOCK, J. H., & VON DER LIPPE, A. Antecedents of optimal psychological adjustment. *J. Consult. & Clin. Psychol.*, 1970, **35**, 283-289.
20. SMITH, M. B. "Mental health" reconsidered: A special case of the problem of values in psychology. *Amer. Psychol.*, 1961, **16**, 299-306.
21. SZASZ, T. S. The uses of naming and the origin of the myth of mental illness. *Amer. Psychol.*, 1961, **16**, 59-65.
22. THOMAS, M., & SEEMAN, J. Personality integration and cognitive processes. *J. Personal. & Soc. Psychol.*, 1972, **24**, 154-161.
23. WHITE, R. W. Motivation reconsidered: The concept of competence. *Psychol. Rev.*, 1959, **66**, 297-333.
24. WINBORN, B. B., & ROWE, W. Self-actualization and the communication of facilitative conditions—A replication. *J. Counsel. Psychol.*, 1972, **19**, 26-29.
25. WRIGHT, L. Components of positive mental health. *J. Consult. & Clin. Psychol.*, 1971, **36**, 277-280.

*Psychology Service*  
*Veterans Administration Center*  
*Hampton, Virginia 23667*



## SOME INTERNATIONAL COMPARISONS OF CANCER MORTALITY RATES AND PERSONALITY: A BRIEF NOTE\*

*College of Education, Aberdeen, Scotland*

---

GORDON RAE AND JAMES MCCALL

---

A few studies have been conducted in Britain and Sweden which tentatively suggest a positive relationship between proneness to cancer and extraversion<sup>1,2,3</sup> and a negative relationship between cancer and neuroticism.<sup>2,4</sup> These have been summarized by Eysenck.<sup>5</sup> Since cigarette smoking increases in linear fashion with degree of extraversion<sup>5,6,7</sup> it might be thought that extraverts develop lung cancer and other types of cancer more readily simply because they smoke more than do introverts. According to Eysenck<sup>5</sup> (p. 118), however, this is improbable.

The present study was an attempt to extend the findings of these earlier studies by demonstrating that an association between cancer and personality holds internationally. National extraversion and anxiety level in eight advanced countries and statistics of the number of cigarettes smoked per adult per annum in these countries have been provided by Lynn and Hayes.<sup>7,8</sup> Mortality rates per 100,000 of population due to lung cancer (males and females separately) and cancer of cervix (females) were also obtained for each of these countries.<sup>9</sup> Respectively for each country these rates were as follows: U. S. A., 45.0, 8.3, 7.7; U. K., 96.7, 17.9, 10.1; Ireland, 43.5, 10.2, 4.2; New Zealand, 40.6, 5.6, 6.8; Canada, 32.8, 5.5, 6.9; West Germany, 58.9, 9.1, 7.6; Australia, 38.8, 5.5, 6.7; Japan, 12.0, 5.0, 3.4.

---

\* Received in the Editorial Office on June 4, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Hagnell, O. Epidemiology and the aged. *Svenska Lakartidn.*, 1961, 58, 492-499.

<sup>2</sup> Kissen, D. M., & Eysenck, H. J. Personality in male lung cancer patients. *J. Psychosomat. Res.*, 1962, 6, 123-127.

<sup>3</sup> Coppen, A., & Metcalfe, M. Cancer and extraversion. *Brit. Med. J.*, 1963, 6 July, 18-19.

<sup>4</sup> Kissen, D. M. Relationship between lung cancer, cigarette smoking, inhalation, and personality. *Brit. J. Med. Psychol.*, 1964, 37, 203-216.

<sup>5</sup> Eysenck, H. J. *Smoking, Health and Personality*. London: Weidenfeld & Nicolson, 1965.

<sup>6</sup> Eysenck, H. J. Personality and cigarette smoking. *Life Sci.*, 1964, 3(7), 777-790.

<sup>7</sup> Lynn, R., & Hayes, B. Some international comparisons of tobacco consumption and personality. *J. Soc. Psychol.*, 1969, 79, 13-17.

<sup>8</sup> Lynn, R. *Personality and National Character*. Oxford, England: Pergamon, 1971.

<sup>9</sup> World Health Statistics Annual. Geneva, Switzerland: World Health Organization, 1969.

The rank order correlations ( $\rho$ ) between national personality levels and cancer mortality rates were as follows (\* indicates  $p < .05$ ): Extraversion  $\times$  Male Lung Cancer,  $+.66^*$ ; Extraversion  $\times$  Female Lung Cancer,  $+.72^*$ ; Extraversion  $\times$  Cancer of Cervix,  $+.64^*$ ; Anxiety  $\times$  Male Lung Cancer,  $-.52$ ; Anxiety  $\times$  Female Lung Cancer,  $-.71^*$ ; Anxiety  $\times$  Cancer of Cervix,  $-.30$ .

The value of  $\rho$  for Cigarette Consumption  $\times$  Lung Cancer (males and females combined) was  $+.07$ ; for Cigarette Consumption  $\times$  Cancer of Cervix,  $+.45$ .

The positive correlations obtained between extraversion and lung cancer rates in both males and females and the negative correlations between these rates and anxiety lend tentative support, on a cross-national basis, to Eysenck's postulate that "persons constitutionally predisposed to take up smoking are also constitutionally predisposed to develop cancer." The correlations found between the two personality variables and cervical cancer are similar to those obtained in the case of lung cancer. Again this could be regarded as evidence linking cancer with constitutional factors, or one might postulate the existence of a mediating agency. For example, Susser and Watson<sup>10</sup> suggest that cervical cancer is related to genital hygiene and to frequency of intercourse, and Eysenck<sup>11</sup> presents evidence that extraverts engage more frequently in intercourse than introverts. Neurotics, however, show a slight tendency to engage less frequently in intercourse than stable persons.

*Department of Educational Psychology*  
*Aberdeen College of Education*  
*Hilton Place*  
*Aberdeen AB9 1FA, Scotland*

<sup>10</sup> Susser, M. W., & Watson, W. *Sociology in Medicine*. London: Oxford Univ. Press, 1967.

<sup>11</sup> Eysenck, H. J. *Psychology Is About People*. London: Lane, 1972.

## RANDOMIZATION TESTS: COMPUTER TIME REQUIREMENTS\*<sup>1</sup>

*The University of Calgary, Canada*

---

EUGENE S. EDGINGTON AND ALLAN R. STRAIN

---

### SUMMARY

Even though randomization tests are the most powerful of nonparametric tests and are the only valid tests to employ when there has been random assignment, but not random selection, of subjects in experiments (a common practice in psychology), such tests are rarely used by psychologists. The limited adoption of randomization tests is primarily a consequence of the great amount of computation they require. The present study shows, however, that the computation for randomization test counterparts of the *t* test and one-way analysis of variance can be relatively inexpensive when performed by a high-speed computer.

---

### A. INTRODUCTION

Moses (8) and Siegel (10) made psychologists aware of nonparametric statistical tests, and at the present time there are many psychologists using nonparametric tests. But practically all of the tests used are rank-order tests. Although both Moses and Siegel described randomization tests, very few psychologists use them. Yet randomization tests are the most powerful of nonparametric tests. Like rank-order tests, randomization tests involve no assumptions about population parameters, but randomization tests utilize raw measurements instead of ranks. Aside from their desirable distribution-free properties, randomization tests have another property that is of considerable importance to psychologists: when subjects have *not* been randomly selected but have been randomly assigned to treatments, randomization tests are the *only* valid tests that can be performed. (Rank-order tests would also be valid, but in fact can be regarded as randomization tests for ranks.) For such experiments parametric tests, being based on random-sampling models, are not valid (5). Random assign-

---

\* Received in the Editorial Office on June 5, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Financial support for this research was given by The University of Calgary.



ment without random sampling is common enough in psychology to make this consideration an important one in assessing the value of randomization tests. Cotton (2, pp. 64-65), after referring to the random-sampling assumption in parametric tests, stated that "Unfortunately, in psychology and probably in most other fields of experimentation, the typical procedure is *randomization*, not *random sampling*" (italics by Cotton). He then went on to explain that it is, therefore, randomization theory, upon which randomization tests are based, not random-sampling theory, that is relevant to most applications of hypothesis testing in psychology.

Thus, there are two major reasons for the importance of randomization tests to psychologists: (a) they are the most powerful of nonparametric tests, and as such many psychologists would find them valuable substitutes for rank-order tests, and (b) they are the *only* completely valid tests to be used when an experimenter has used randomization on nonrandomly selected subjects, which is common practice in psychology.

#### B. THE FEASIBILITY OF RANDOMIZATION TESTS

In view of the advantages of randomization tests it may seem incredible that very few psychologists use them for testing hypotheses. One reason for their infrequent application is that there is little literature on the application of randomization tests. Fisher (6) and Pitman (9) developed randomization tests in the 1930s, but much of the literature on randomization tests since that time has concerned theory rather than application. Kempthorne (7) is perhaps the major contributor to randomization test theory although he has discussed application as well. Randomization tests are seldom discussed or even mentioned in statistics books; consequently, many psychologists are only vaguely aware of their existence. There are, however, two books on distribution-free statistics, written by psychologists, which discuss randomization tests at length (1, 3).

There is another, more pervasive, reason for randomization tests being rarely used by psychologists: the great amount of computation that randomization tests require. This has been in fact the only serious objection that has been raised against the widespread employment of randomization tests as alternatives to more conventional tests, like the  $t$  test or analysis of variance. For example, an independent  $t$  test with only seven subjects in each of the groups requires  $14!/7!7! = 3432$  randomizations or divisions of the 14 scores in every way into two groups of seven scores each, and for each division  $t$  or some other test statistic, like the difference between

means, must be computed. (As will be shown in Section C, only half of these need to be performed, but that is still a large number.)

Thirty or 40 years ago such an objection was reasonable, but the accessibility of high-speed computers in modern times renders it ineffectual. As will be shown in Section C, it is inexpensive in terms of both time and money to run a randomization test involving several thousand randomizations. And where the number of randomizations for an exact randomization test is too large for even a computer to handle economically, an approximate randomization test utilizing a random sample of all randomizations can be performed. To obtain such a random sample, a pseudo-random number generator is used to partition the data randomly a specified number of times. Each of the approximate randomization tests in the present study used a random sample of 999 randomizations. Under the null hypothesis, the observed set of data is also a randomly selected randomization, and so under the null hypothesis it along with the other 999 randomizations provides a random sample of 1000 randomizations. The rationale for the approximate randomization test has been presented elsewhere (3, p. 152-155; 4). An approximate randomization test using 999 randomizations is a perfectly valid test in its own right and is almost as powerful as an exact randomization test no matter how many randomizations the exact test requires. The power of approximate randomization tests will be discussed in Section C.

### C. EXAMPLES OF COMPUTER TIME REQUIREMENTS

Computer programs were developed for exact and approximate randomization tests for four statistical tests: (a) independent  $t$  test, (b) correlated  $t$  test, (c) simple one-way analysis of variance, and (d) repeated-measures analysis of variance.

The computer programs were used on a CDC 6400 computer for analysis of the data for various sample sizes and, in the case of analysis of variance, for various numbers of groups. Table 1 shows the results of computer trials for the various tests. The time shown in Table 1 is the total computer time. The cost shown includes the cost of both computer processing and computer memory, but does not include the cost of reading the punched cards, which would generally add about 15¢ to the cost given in the table. For each computer trial, a probability value for the observed results was determined. The probability values, however, are not shown because they had no influence on the computer time or cost; the time taken to compare every test

TABLE 1  
EXAMPLES OF COMPUTER TIME AND COST FOR EXACT AND  
APPROXIMATE RANDOMIZATION TESTS

Type of test	Number of groups	Number of Ss per group	Time (seconds)	Cost (dollars)
Exact $t$	2	8	7.1	\$ .42
Exact $t$	2	9	26.4	1.52
Exact $t$	2	10	118.8	6.78
Approximate $t$	2	10	2.0	.12
Approximate $t$	2	50	9.1	.52
Exact correlated $t$	2	10	1.1	.08
Exact correlated $t$	2	13	7.4	.50
Exact correlated $t$	2	15	29.1	1.95
Approx. correlated $t$	2	10	2.3	.15
Approx. correlated $t$	2	50	9.7	.62
Exact simple ANOVA	3	3	.5	.04
Exact simple ANOVA	3	4	6.2	.37
Exact simple ANOVA	4	3	18.0	1.04
Approx. simple ANOVA	3	4	1.9	.12
Approx. simple ANOVA	5	10	9.6	.55
Approx. simple ANOVA	5	20	18.7	1.07
Exact repeated-measures	3	3	.3	.03
Exact repeated-measures	3	5	2.0	.14
Exact repeated-measures	4	4	24.3	1.63
Approx. repeated-measures	3	10	4.8	.31
Approx. repeated-measures	4	15	12.0	.77
Approx. repeated-measures	5	10	10.9	.70
Approx. repeated-measures	10	10	30.0	1.92

statistic with the observed test statistic is independent of the proportion of those test statistics that are as large as the observed value.

Table 1 shows that exact randomization tests are inexpensive for small samples. An exact independent  $t$  test costs only a few dollars for 10 subjects per group. And an exact correlated  $t$  test for 15 pairs of measurements is inexpensive as well. It must be stressed, however, that the computer time and cost for these tests increase rapidly with sample size. Increasing the number of subjects per group from 10 to 11 for the exact independent  $t$  test would approximately *quadruple* the computer time. (The number of randomizations is not quite quadrupled, but the increase in the amount of data to be handled with each randomization along with the increase in the number of randomizations would approximately quadruple the computer time.) And each additional increase of one subject per group would approximately quadruple the computer time and cost. Thus, increasing the number of subjects per group from 10 to 12 would increase the cost from \$6.78 to about \$100. For the exact correlated  $t$  test, adding one more pair of measurements doubles the number of randomizations, and so the computer time approximately doubles.

Although the rate of increase in cost is less than that for the exact independent  $t$  test, it is rapid enough to restrict the utility of the exact correlated  $t$  test to samples of no more than 17 or 18 pairs of measurements.

The exact randomization test for simple one-way analysis of variance test is inexpensive only for rather small samples. With three groups and four measurements per group, the computer time is slightly over six seconds, but increasing the number of measurements per group to five would multiply the number of randomizations (and consequently the computing time) by more than 20. With four groups, the effect of a slight increase in the number per group is even greater. The table shows that it takes 18 seconds for analyzing the data from four groups with three subjects per group, but to increase the number of subjects to four per group is out of the question because this would lead to about 170 times as many randomizations.

For the exact repeated-measures test, with three groups, adding one more measurement per group by adding another subject multiplies the number of randomizations by six. Thus, although it takes only two seconds for three groups with five subjects per group, it would take about 70 seconds to perform an exact test with seven subjects per group. With four groups, adding one more measurement per group multiplies the number of randomizations by 24, so four measurements per group is about the largest practical sample size for the exact repeated-measures test.

Thus, we see that for the four types of test under consideration, electronic computer performance of exact randomization tests is inexpensive for small samples, but not for moderately large ones.

Approximate randomization tests with 999 randomizations, on the other hand, are seen in Table 1 to be inexpensive even for large samples, for all four kinds of tests. The power of an approximate randomization test with 999 randomizations is almost that of the exact randomization test, no matter how many randomizations the exact test requires. For example, for data where the exact test would provide a probability of .01, the probability is .99 that the approximate test would give a probability value no greater than .018, and where the exact test probability would be .05, the probability is .99 that the approximate test probability would be no greater than .067 (3, pp. 152-155; 4). Increasing the number of randomizations used in an approximate test would, of course, increase the power of the approximate test, and the computer time would be proportional to the number of randomizations because there would simply be more operations of the same kind. Again, it should be kept in mind that just as a rank-order test is a valid test even



though it does not have the power of an exact randomization test, so also is an approximate randomization test a *valid* test even though its *power* is not quite that of the exact test.

#### D. COMPUTATIONAL SHORTCUTS

The computer times and costs given in Table 1 are conditional on certain shortcuts that minimized the computation. A shortcut that concerned all of the tests was to use  $\Sigma T^2$ —that is, the sum of the squares of the group totals—as a test statistic instead of  $t$  or  $F$ . For a randomization test, two test statistics are equivalent (in the sense of providing the same probability values) if one is a monotone function of the other. Since  $t^2 = F$  for both independent group comparisons and comparisons of correlated groups,  $t$  and  $F$  are equivalent test statistics. The test statistic  $\Sigma T^2$  is a monotone function of  $F$  for both simple one-way analysis of variance (with equal  $n$ s) and for repeated-measures analysis of variance (1, pp. 81-83), and so it is an equivalent test statistic to both  $F$  and  $t$  for the tests involved in this study. The use of  $\Sigma T^2$  as a test statistic gives the probability that would be obtained if  $F$  or  $t$  were used and saves considerable time for randomization tests, where thousands of computations of the test statistic may be made for performing a single test.

The other principal computational shortcut affected only the *exact* randomization tests. It consisted in reducing the number of randomizations that were required. For example, the independent  $t$  test with equal  $n$ s requires only half of the  $(n_1 + n_2)!/(n_1)!(n_2)!$  randomizations to be performed because of the symmetry of the sampling distribution of the test statistic  $\Sigma T^2$ . For every randomization of scores between the treatments, there is a corresponding mirror-image randomization where the scores are divided in the same way, but where the group designations (like "Group 1" or "Group 2") are reversed. For example, if for one of the randomizations of six scores into three scores for each group, Group 1 has the scores 6, 8, and 12, and Group 2 has the scores 5, 9, and 17, then there must be another randomization where Group 2 has the scores 6, 8, and 12, and Group 1 has the scores 5, 9, and 17. Since the test statistic  $\Sigma T^2$  is the same for one randomization as for the mirror-image randomization, only one member of each such pair of randomizations need be used to work up the exact sampling distribution of the test statistic; the other member is redundant.

A similar but greater reduction in the number of randomizations to be computed was made for the exact randomization test for simple one-way analysis of variance. Once again this reduction was possible because of



equal sample sizes in the various groups and could not be made with unequal sample sizes. As was explained above, the number of randomizations for the exact  $t$  test could be cut in half because there were two ways of assigning the two group designations to any particular split of scores. With three groups, in analysis of variance, there are  $3! = 6$  different ways of assigning the labels "Group 1," "Group 2," and "Group 3" to a particular division of scores into three groups, and so only  $1/6$ th of the randomizations need to be computed, the others being redundant because they would provide the same test statistic value. For simple one-way analysis of variance, then, only  $1/G!$  of the randomizations needed to be computed, where  $G$  is the number of groups.

The same considerations reduced the number of required randomizations for the exact tests for correlated  $t$  and repeated-measures analysis of variance to  $1/G!$  of the total number possible.

#### REFERENCES

1. BRADLEY, J. V. *Distribution-Free Statistical Tests*. Englewood Cliffs, N. J.: Prentice-Hall, 1968. Pp. 81-83.
2. COTTON, J. W. *Elementary Statistical Theory for Behavior Scientists*. Reading, Mass.: Addison-Wesley, 1967. Pp. 64-65.
3. EDGINGTON, E. S. *Statistical Inference: The Distribution-Free Approach*. New York: McGraw-Hill, 1969.
4. ———. Approximate randomization tests. *J. of Psychol.*, 1969, **72**, 143-149.
5. ———. The random-sampling assumption in "Comment on component-randomization tests," *Psychol. Bull.*, in press.
6. FISHER, R. A. *The Design of Experiments* (6th ed.). New York: Hafner, 1951.
7. KEMPTHORNE, O. The randomization theory of experimental inference. *J. Amer. Stat. Assoc.*, 1955, **50**, 946-967.
8. MOSES, L. E. Nonparametric statistics for psychological research. *Psychol. Bull.*, 1952, **49**, 122-143.
9. PITMAN, E. J. G. Significance tests which may be applied to samples from any populations. *J. Royal Stat. Soc. (Series B)*, 1937, **4**, 119-130.
10. SIEGEL, S. *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill, 1956.

*Department of Psychology*  
*The University of Calgary*  
*2920-24th Avenue N. W.*  
*Calgary, Alberta, Canada T2N 1N4*



## CATHOLIC SISTERS AND THE EDWARDS PERSONAL PREFERENCE SCHEDULE\*

*College of Saint Teresa*

---

HARRY W. GARDINER

---

### SUMMARY

The Edwards Personal Preference Schedule was administered to a group of 47 Roman Catholic nuns ranging in age from 22 to 55. Results were compared with normative data for female college students and women in the general adult population. Catholic sisters were shown to be significantly higher than college women in the need for deference, abasement, nurturance, and endurance and lower in exhibition, dominance, change, and heterosexuality. Compared with adult women, they were higher in intraception and aggression and lower in deference, order, and endurance. Catholic sisters were significantly higher than both groups in affiliation and succorance and lower in achievement.

---

### A. INTRODUCTION

A number of previous investigations have attempted to predict the presence or absence of various personality characteristics on the basis of knowledge of the religious value systems practiced by selected individuals (1, 3, 4, 6). The present study represents one of a series of small-scale investigations aimed at understanding the specific processes involved in personality development within a community of Roman Catholic sisters.

### B. METHOD

The Edwards Personal Preference Schedule (2), a forced-choice inventory developed to assess a large number of the psychological needs presented by Murray (5), was administered to 47 Roman Catholic nuns enrolled in courses at a private liberal arts college during the summer of 1969. Subjects ranged in age from 22 to 55 with a mean age of 35 years.

---

\* Received in the Editorial Office on June 20, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

## C. RESULTS

Means and standard deviations for each of the 15 personality variables measured by this inventory were obtained and are presented in Table 1.

Comparisons were made, by means of two-tailed *t* tests, with normative data for two groups of subjects: female college students and women in the general adult population (2).

TABLE 1  
MEANS AND STANDARD DEVIATIONS OF THE EPPS VARIABLES FOR  
ROMAN CATHOLIC SISTERS AND TWO FEMALE NORM GROUPS<sup>a</sup>

Variable	Catholic sisters		College women		Adult women	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Achievement	10.94	3.97	13.08	4.19	13.58	3.95
Deference	13.61	3.48	12.40	3.72	14.72	3.84
Order	11.36	3.80	10.24	4.37	15.59	4.57
Exhibition	11.94	3.61	14.28	3.65	11.48	3.88
Autonomy	12.06	3.35	12.29	4.34	12.10	4.11
Affiliation	20.60	3.90	17.40	4.07	17.76	4.15
Intracception	16.77	4.45	17.32	4.70	15.28	4.13
Succorance	16.89	4.34	12.53	4.42	12.86	4.55
Dominance	11.02	5.25	14.18	4.60	10.24	4.73
Abasement	16.91	4.60	15.11	4.94	16.89	4.88
Nurturance	19.31	3.39	16.42	4.41	18.48	4.43
Change	15.79	4.51	17.20	4.87	15.99	4.73
Endurance	14.09	4.81	12.63	5.19	16.50	4.66
Heterosexuality	6.79	5.37	14.34	5.39	8.12	6.59
Aggression	11.55	4.84	10.59	4.61	10.16	4.37

<sup>a</sup> Figures for two norm groups reproduced by permission. Copyright 1954, © 1959 by The Psychological Corporation, New York, New York. All rights reserved.

Statistical analysis revealed a number of interesting differences between the obtained data and those reported in the manual. While the normative data in the manual are nearly two decades old and in need of revision, norms of more recent date are not available for comparison.<sup>1</sup> Therefore, the reader should consider the possibility that the results reported here may reflect either real differences between the groups or the effects of time on all groups. Only additional research can more adequately address itself to this point.

Scores indicated that when compared with female college students, Catholic sisters showed a significantly higher need for deference, order, affiliation, succorance, abasement, nurturance, and endurance ( $t = 3.09, 2.61, 7.29, 8.92, 3.47, 7.57$ , and  $2.69$ , respectively;  $p < .01, .01, .001, .001, .001, .001$ , and  $.01$ , respectively). They were significantly lower in the need for achieve-

<sup>1</sup> Personal communication, Test Division Research Section, Psychological Corporation.

ment, exhibition, dominance, change, and heterosexuality ( $t = 4.79, 5.76, 5.34, 2.77, \text{ and } 12.49$ ;  $p < .001, .001, .001, .01, \text{ and } .001$ ). There were no significant differences between the groups on the three remaining variables: autonomy, intraception, and aggression.

When compared with women in the adult population, Catholic sisters were significantly higher on the following needs: affiliation, intraception, succorance, nurturance, and aggression ( $t = 6.47, 2.97, 8.25, 2.17, \text{ and } 2.55$ ;  $p < .001, .01, .001, .05, \text{ and } .05$ ). Significantly lower scores were found among sisters on achievement, deference, order, endurance, and heterosexuality ( $t = 5.91, 2.83, 9.89, 4.45, \text{ and } 2.20$ ;  $p < .001, .01, .001, .001, \text{ and } .05$ ). The groups did not differ significantly on exhibition, autonomy, dominance, abasement, and change.

#### D. CONCLUSIONS

These results yield a profile of a Catholic sister which tends to support the general perception held by many lay people: i.e., one who is generally high on affiliation, with feelings that personal pain and misery suffered do more good than harm, and ready to assist others who are less fortunate, while at the same time being low on the need for exhibition, desire to be a leader in groups to which she belongs, and the need for heterosexuality. Among the more interesting findings are those that indicate sisters to be significantly lower in achievement than women in college and in the general adult population, higher in deference than college women but lower than adult women, higher in endurance than college women but lower than those in the adult population, and higher in aggression than adult women.

Also of interest, but more difficult to explain, are those areas in which similarities occurred. For example, sisters did not differ from either college or adult women on autonomy. It may be that recent changes within religious orders have greatly increased the opportunities for nuns to express themselves in less conventional ways, avoid group conformity, and more openly criticize those in authority positions. On the other hand, why sisters and adult women should score so much alike on such a variable as abasement, indicating a need to feel guilty when one does something wrong, as well as a need for punishment, along with a feeling of depression when unable to handle situations, remains a surprising finding. This result, as well as others, suggests areas of personality development within members of a specific religious community which are deserving of further investigation.



## REFERENCE

1. CARMICHAEL, J. J. Some relationships of denominational affiliation to selected personal characteristics of male college students. Unpublished Doctoral dissertation, Ohio University, Athens, 1963.
2. EDWARDS, A. L. Manual for the Personal Preference Schedule. New York: Psychol. Corp., 1954.
3. GARDINER, H. W. Philosophies of human nature among Roman Catholic sisters. *Psychol. Rep.*, 1972, **30**, 369-370.
4. HJELLE, L. A., & ABOUD, J., JR. Some personality differences between seminarians and nonseminarians. *J. Soc. Psychol.*, 1970, **82**, 279-280.
5. MURRAY, H. A. Explorations in Personality. New York: Oxford Univ. Press, 1953.
6. TENNISON, J. C., & SNYDER, W. U. Some relationships between attitudes toward the church and certain personality characteristics. *J. Counsel. Psychol.*, 1968, **15**, 187-189.

*Department of Psychology*  
*College of Saint Teresa*  
*Winona, Minnesota 55987*

## PERSONALITY AND DEMOGRAPHIC CORRELATES OF SIMULATED PERSONAL SPACE\*<sup>1</sup>

Brigham Young University

DARHL M. PEDERSEN<sup>2</sup>

### SUMMARY

Personality, demographic, and simulated personal space measures were administered to 170 male college students. Correlations of the personality and demographic variables with the personal space measures indicated that Ss with a smaller personal space tended to be less aggressive, be more tolerant of ambiguity, be more self-acceptant, and have a higher ideal self. They also tended to be of Southern European extraction. Ss who had a smaller *back* personal space tended to have higher self-acceptance, higher emotional stability, and a lower consumption of alcohol.

### A. INTRODUCTION

Personal space has been defined as "the area immediately surrounding the individual in which the majority of his interactions with others take place" (9, p. 237). Measures of personal space have been classified as behavioral and simulated (15). Measures that utilize live interaction between Ss are behavioral. Measures in which Ss manipulate representations of themselves and others are simulated. Size of personal space varies from person to person both between cultures and within a culture. Watson and Graves (19) found that Arabs interacted closer in a behavioral task than did Americans. Using a simulated task to study five national groups, Little (10) found that Greeks, Americans, and Italians had the closest personal space followed by Swedes followed by Scots.

There has been little success in finding determinants of personal space in our culture, even though large individual differences occur. However, a

\* Received in the Editorial Office on June 25, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was completed while the author was a National Research Council Postdoctoral Resident Research Associate at NASA's Ames Research Center and was supported in part by a Faculty Research Fellowship granted by Brigham Young University.

<sup>2</sup> Requests for reprints should be sent to the author at the address shown at the end of this article.

larger personal space has been found for schizophrenics (7, 17), introverts (12, 20), and males (11, 14). Hornbrook and Tateishi (6) found no relationship between the California Personality Inventory and simulated personal space involving the placement of dolls. The purpose of the present study was to find some personality and demographic correlates of simulated personal space.

## B. METHOD

### 1. Subjects

Ss were 170 male students from junior colleges in the Bay Area of California. They received \$2.50 per hour to participate in the experiment.

### 2. Measuring Instruments

a. *Psychological Audit for Interpersonal Relations: The "Pair" test.* The Pair test consists of 500 true-false statements that measure attributes important in interpersonal relations (18). Two scales are scored: aggressive hostility and self-acceptance.

b. *Cycloid Disposition Scale.* On the basis of a factor analysis, Becker (1) has determined that there is considerable overlap among the factors measured by the Guilford (5) and Cattell (3) inventories. The most important factor in the inventories is "anxiety-emotional stability." The Cycloid Disposition Scale of Guilford (5) loads highly on that factor; therefore, it was selected to represent that factor. It measures neurotic tendency or emotional maladjustment.

c. *Tolerance of Ambiguity Scale.* This Likert-type scale consists of 16 items (2). Eight are phrased so that agreement indicates tolerance of ambiguity, and eight are phrased so that agreement reflects intolerance of ambiguity. This format reduces the effects of the acquiescence response set. Alpha-coefficient reliability estimates from .39 to .62 and test-retest reliability estimates of .85 have been reported. The items are scored so that the overall score indicates *tolerance* of ambiguity.

d. *Adjective Check List.* This measure consists of 300 adjectives (4). Ss checked those adjectives that are self-descriptive. Seventy-five of the adjectives have been classified as "most favorable." A Self-Acceptance score is obtained by dividing the number of most favorable adjectives checked by the total number of adjectives checked and expressing the quotient to the nearest percentage.

e. *Self and Others Rating Scale.* This semantic differential is constructed by use of 25 bipolar adjectives which have high loadings on the evaluative

(i.e., good-bad) dimension of meaning (13). Each pair of adjectives is separated by an eight-point scale. Ratings are obtained for Self Concept (yourself as you really are) and Ideal Self (the kind of person you would like to be). The favorable adjective in each pair is randomly assigned to the right or left side. Total scores are found by scoring each bipolar adjective scale from 1 to 8 (with an 8 for the scale closest to the favorable adjective) and summing these scores over all 25 scales. Split-half reliability coefficients of .92 for Self Concept and .85 for Ideal Self have been obtained (13).

f. *Body-Accessibility Questionnaire*. In this measure developed by Jourard (8) Ss use a diagram of 24 body regions to indicate the areas they have touched and those on which they have been touched relative to their best male friend. An odd-even reliability estimate of .98 has been reported for tactual interchange between male Ss and their same-sexed friends. "Touch" and "touched by" scores are obtained by counting the number of regions checked.

g. *Demographic data questionnaire*. A questionnaire was administered which obtained information concerning (a) birth order (oldest child = 2, not oldest = 1); (b) number of beers and/or cocktails drunk per day; (c) self-rating of body attractiveness on a four-point scale (low = 1, high = 4); and (d) Southern European extraction (yes = 2, no = 1).

h. *Pedersen Personal Space Measure*. This simulated measure of personal space required Ss to place a mobile profile of a man relative to a stationary profile of a man printed on a page (15). The profiles used are shown in Figure 1. The profiles are facing left (L), facing front (F), facing right (R), and top (T). Ss are instructed to make the distance between the profiles representing themselves and another male approximately the same age "as close as is comfortable for you in most social situations." There are 20 pairs of profiles, one pair per page. Except that "self" is printed under the left profile for the first 10, and "other" is printed under the left profile for the second 10, the two sets are the same. The profiles for each set are as follows: R-L, F-L, L-L, R-F, F-F, L-F, R-R, F-R, L-R, T-T. The score for each item is the distance between profiles to the nearest millimeter. Test-retest reliabilities ranging from .71 to .91 have been reported for the item scores (15). The total score is the sum of the item scores.

i. *Rawls Personal Space Measure*. Figure drawings of five views of a man are presented, one view per page (16). The views (shown in Figure 2) are facing front, facing left, facing back, facing right, and top. Ss are instructed to "draw a circle around the figure showing how far you would like people to stay from you during ordinary social situations." The distances at the

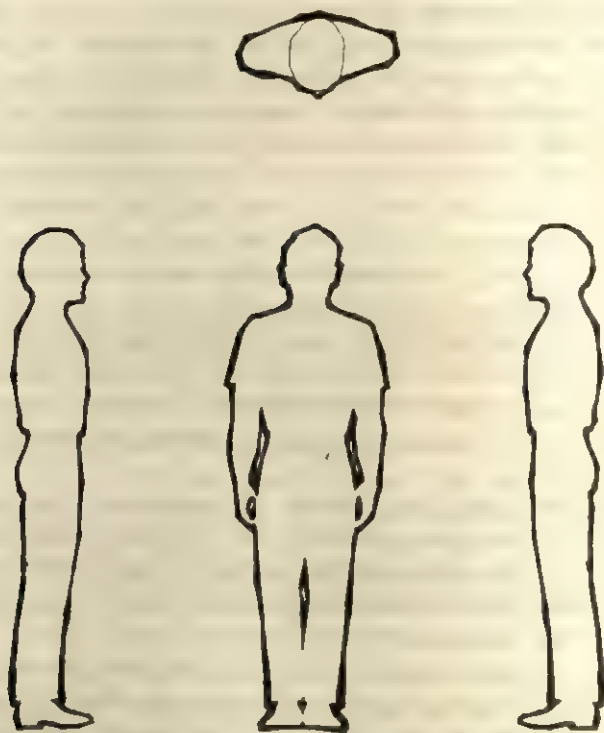


FIGURE 1  
PROFILES USED IN THE PEDERSEN PERSONAL SPACE MEASURE

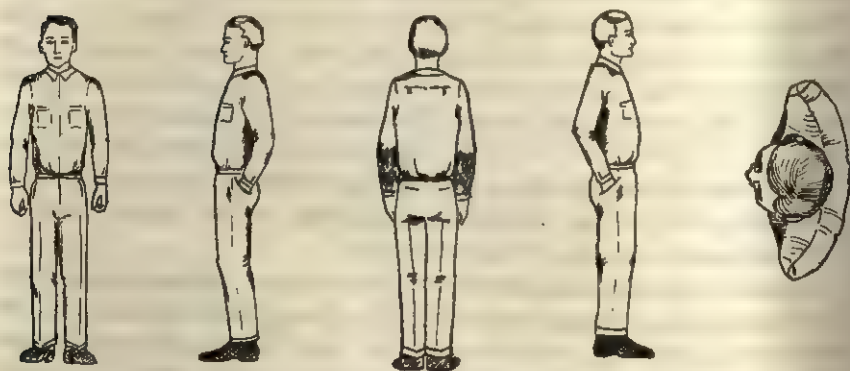


FIGURE 2  
FIGURE DRAWINGS USED IN THE RAWLS PERSONAL SPACE MEASURE



belt line from the front of the man to the perimeter of the circle in front and from the back of the man to the perimeter of the circle in back are obtained to the nearest millimeter for three profiles: facing left, facing right, and top. Two scores are obtained: Rawls: Front, the mean of the three front distances, and Rawls: Back, the mean of the three back distances. The test-retest reliabilities have been reported as .90 for both the Rawls: Front and Rawls: Back scores.

### 3. Procedure

The results of previous studies did not encourage the use of a single personality test. Therefore, personality and demographic measures were selected to cover a broad range of such variables. The selection included variables important to interpersonal interaction; neuroticism; tolerance of ambiguity; ratings of self (including self-acceptance, self-concept, self-ideal, and body attractiveness); amount of touching relative to best male friend; birth order; cultural extraction; and amount of drinking. Ss were administered the personality and demographic measures in an initial session. They were given the personal space measures in a session approximately two weeks later.

### C. RESULTS AND DISCUSSION

Means, standard deviations, and correlations with the simulated personal space variables for the personality and demographic variables are given in Table 1. Three comparisons involving means were meaningful. The rating of Self Concept compared to that for Ideal Self confirms that the Ss in general would like to be better than they are. The number of body regions touched by best male friend was approximately equal to the number of his regions touched. Finally, the mean distance to the back of the Rawls Personal Space Measure was slightly larger than the front distance, indicating that Ss tended to prefer people to be a little more distant at the back as compared to the front.

None of the correlations between the variables and the personal space measures was very large. However, all but three of the variables were significantly correlated with at least one personal space variable, and two of those three barely missed significance. For this type of data correlations are characteristically not large. They do indicate some encouraging trends. The patterns of correlations with the Pedersen and Rawls measures were somewhat distinct, indicating that the two measures were not highly correlated. The Pedersen measure correlated .36 with the Rawls: Front score and .39

TABLE 1  
MEANS AND STANDARD DEVIATIONS OF PERSONALITY AND DEMOGRAPHIC VARIABLES AND  
CORRELATIONS WITH PERSONAL SPACE

Variable	Mean	SD	Pedersen: Total	Correlations Rawls: Front	Rawls: Back
Pair test					
Aggressive Hostility	11.51	3.92	.15*	.13	.14
Self-Acceptance	14.79	4.32	-.07	-.11	-.16*
Cycloid Disposition Scale	19.65	9.20	.04	.13	.15*
Tolerance of Ambiguity Scale	15.14	10.51	-.16*	-.12	-.12
Adjective Check List					
Self-Acceptance	43.29	11.03	-.16*	-.18*	-.09
Self and Other Rating Scale					
Self Concept	149.88	17.74	-.05	-.06	-.04
Ideal Self	175.93	20.33	-.18*	-.10	-.11
Body-Accessibility Questionnaire					
Touched by best male friend	23.45	6.36	-.05	-.18*	.01
Touch best male friend	23.42	6.36	-.06	-.17*	-.01
Demographic data questionnaire					
Oldest child	1.34	.49	.04	-.14	-.06
Number of beers and/or cocktails per day	.45	1.14	.20**	.03	.15*
Body attractiveness	3.08	.69	.09	.14	.14
Southern European extraction	1.26	.44	-.15*	-.10	-.09
Pedersen Personal Space Measure: Total	30.22	12.45		.36**	.39**
Rawls Personal Space Measure					
Front	21.10	10.19			.69**
Back	22.60	14.44			

\*  $p \leq .05$  (two-tailed).

\*\*  $p \leq .01$  (two-tailed).

with the Rawls: Back score. Also, the Front and Back scores obtained from the Rawls measure were quite different ( $r = .69$ ).

The pattern of correlations with the Pedersen measure indicated that Ss with smaller personal space tended to have less aggressive hostility, greater tolerance of ambiguity, greater self-acceptance, and higher ideal self; drink fewer beers and/or cocktails per day; and be of Southern European extraction. These characteristics form a personality syndrome of the person with a small personal space. The correlations with the Rawls: Front score added to the syndrome greater body-accessibility with best male friend. A picture emerged of a person who was nonaggressive in interpersonal relations and was comfortable with himself. He not only tolerated ambiguous situations, but he also accepted himself. He did not have a higher self-concept, but he did want to be better than he was. The fact that he drank less may be tied to his acceptance of himself and others. It may be that people who drink more do so to some extent because of personal and social adjustment problems. The fact that small personal spacers were of Southern European extraction was consistent with research findings that Southern Europeans tend to have smaller personal space (10). The multiple correlation of number of beers and/or cocktails, ideal self, and tolerance of ambiguity with the Pedersen Personal Space Measure was .32. And the multiple correlation of self-acceptance and being touched by best male friend with the Rawls: Front score was .26. The correlations with the Rawls: Back score indicated the type of person that did not like others to approach closely from behind. People who had smaller *back* personal space tended to have higher self-acceptance, have higher emotional stability, and drink less. Together these characteristics depicted a person who was relatively secure.

Even though the magnitudes of the correlations obtained were not large, they presented a consistent picture of the type of person who had a small personal space in general and a small personal space behind him. Further research using variables that fill in the syndromes described should enhance an understanding of the personality determinants of personal space.

#### REFERENCES

1. BECKER, W. C. A comparison of the factor structure and other properties of the 16 PF and the Guilford-Martin Inventories. *Educ. & Psychol. Meas.*, 1961, **21**, 393-404.
2. BUDNER, S. Intolerance of ambiguity as a personality variable. *J. Personal.*, 1962, **30**, 29-50.
3. CATTELL, R. B., SAUNDERS, D. R., & STICE, G. Handbook for the 16 Personality Factor Questionnaire. Champaign, Ill.: Inst. Personal. & Ability Test., 1957.

4. GOUGH, H. G. The adjective check list as a personality research technique. *Psychol. Rep.* (Monog. Suppl. 2), 1960, **6**, 107-122.
5. GUILFORD, J. P. An Inventory of Factors STDCR. Beverly Hills, Calif.: Sheridan Supply, 1940.
6. HORN BROOK, M. C., & TATEISHI, P. M. Personality correlates of judged appropriate interaction distances. Unpublished manuscript, University of Denver, Denver, Colorado, undated.
7. HOROWITZ, M. J., DUFF, D. F., & STRATTON, L. O. Body-buffer zone. *Arch. Gen. Psychiat.*, 1964, **11**, 651-656.
8. JOURARD, S. M. An exploratory study of body-accessibility. *Brit. J. Soc. & Clin. Psychol.*, 1966, **5**, 221-231.
9. LITTLE, K. B. Personal space. *J. Exper. Soc. Psychol.*, 1965, **1**, 237-247.
10. ———. Cultural variations in social schemata. *J. Personal. & Soc. Psychol.*, 1968, **10**, 1-7.
11. LONG, B. H., ZILLER, R. C., & HENDERSON, E. H. Developmental changes in the self-concept during adolescence. *Sch. Rev.*, 1968, **76**, 210-230.
12. PATTERSON, M. L., & HOLMES, D. S. Social interaction correlates of the MMPI extroversion-introversion scale. *Amer. Psychol.*, 1966, **21**, 724-725.
13. PEDERSEN, D. M. Evaluation of self and others and some personality correlates. *J. of Psychol.*, 1969, **71**, 225-244.
14. ———. Developmental trends in personal space. *J. of Psychol.*, 1973, **83**, 3-9.
15. ———. Development of a personal space measure. *Psychol. Rep.*, 1973, **32**, 527-535.
16. RAWLS, J. R., TREGO, R. E., & MCGAFFEY, C. N. A comparison of personal space measures. Report of NASA Grant NGR-44-009-008, Institute of Behavioral Research, Texas Christian University, Fort Worth, October, 1968.
17. SOMMER, R. Studies in personal space. *Sociometry*, 1959, **22**, 247-260.
18. STEPHENSON, R. R. Psychological Audit for Interpersonal Relations: The "Pair" Test. Los Angeles: Author, 1966.
19. WATSON, O. M., & GRAVES, T. D. Quantitative research in proxemic behavior. *Amer. Anthropol.*, 1966, **68**, 971-985.
20. WILLIAMS, J. L. Personal space and its relation to extroversion-introversion. Unpublished Master's thesis, University of Alberta, Edmonton, Canada, 1963. Cited by R. Sommer, *Personal Space*. Englewood Cliffs, N.J.: Prentice-Hall, 1969.

Department of Psychology

1230A SFLC

Brigham Young University

Provo, Utah 84601

## SEMANTIC STYLE VARIANCE IN PERSONALITY QUESTIONNAIRES\* <sup>1</sup>

*University of Rochester*

---

ROBERT STRAHAN<sup>2</sup> AND KATHLEEN CARRESE GERBASI

---

### SUMMARY

Terms such as *often*, *usually*, and *sometimes* frequently modify items found in personality questionnaires. *Semantic style* refers to individual differences in the quantitative meanings subjects give such terms. This study investigated the relation between semantic style variables and questionnaire measures of social desirability, extraversion, and neuroticism. Significant associations were found for two of these three personality scales: social desirability and neuroticism. If semantic style is shown to be pervasive and appreciable in its influence on questionnaire response, its systematic control should become a psychometric goal of some importance. Techniques are discussed for thus achieving more valid questionnaire measures.

---

### A. INTRODUCTION

Terms such as *often*, *very*, *rather*, and *sometimes*, used frequently in everyday speech to give greater precision to one's meaning, are utilized in psychological questionnaires as well for the same purpose. Paradoxically, however, these terms have themselves a considerable degree of ambiguity (5, 7, 8). This fact is, of course, acknowledged by psychometricians (e.g., 4, 6), and subjects as well have been known to remark on the point. Notwithstanding this awareness, little systematic research appears to have been directed toward what might be termed *semantic style* factors in questionnaire response.

This study sought to determine whether measures of individual differences

---

\* Received in the Editorial Office on June 27, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Part of this work was supported by National Institute of Mental Health Grant MH 17677-01, awarded to the senior investigator, and the work was partly undertaken during the junior investigator's tenure as a National Science Foundation Undergraduate Research Participant. Extensive use was made of the facilities of the Computing Center of the University of Rochester, which is partially funded by National Science Foundation Grant GJ-828.

<sup>2</sup> Direct requests for reprints to the first author at the address shown at the end of this article.



in semantic style—specifically, the quantitative meanings given certain modifier terms—are related to personality questionnaires whose items make use of those terms.

## B. METHOD

### 1. Subjects and Procedure

Packets containing a variety of self-report measures were given to about 300 students at the end of an introductory psychology class period. Most of these packets, completed outside of class, were returned to the experimenters a few days later for credit toward a research participation requirement of the course. Contents of the packets differed, with 78 students receiving the materials described below. Forty subjects completed these materials and were the source of data for this study. Twenty of the Ss were males, 20 females.

### 2. Materials

Three self-report personality measures were presented in the intermixed true-false items of a single questionnaire. These were the *Marlowe-Crowne Social Desirability Scale* (MC) of Crowne and Marlowe (2) and *Extraversion* (E) and *Neuroticism* (N) scales of Eysenck and Eysenck (3). Sub-scales of MC were later formed, consisting of the 15 items keyed false or negative (MC Neg) and 15 of the 18 items keyed true or positive (MC Pos).<sup>3</sup> E and N were each composed of 20 items, though one N item was not scored because of its obscurity to American Ss (the E and N measures originated in England).

A *Semantic Scaling Inventory* (SSI) listed 20 frequency quantifiers (e.g., *seldom, usually, occasionally*) that S was to place in what he judged to be their proper order and spacing (ties permitted) on a 5.75-inch vertical line the ends of which were labeled *least possible frequency* and *greatest possible frequency*. Each quantifier term was later scored zero to 100 according to how high (toward the *greater pole*) S located it.

### 3. Rationale of the Analysis

Although the original purpose of the SSI had been to determine equidistant quantifier terms of low variance for use in gradated response questionnaires, the applicability of this inventory to issues of semantic style was recognized. Since a number of SSI terms (or what were judged to be

<sup>3</sup> For purposes incidental to this study, three of the true-keyed items were omitted so that MC Pos and MC Neg would be balanced with respect to scale length.

synonyms) were used by items in the concurrently administered MC, E, and N scales, it was possible to examine the relations of these semantic style terms to the personality measures.

It seemed reasonable that the quantitative meaning an *S* gave to a particular SSI term would have a role in determining whether questionnaire items employing that term were to be endorsed. For example, when presented with the N item, "My mood often goes up and down," *S*'s response should have been influenced in part by the absolute quantitative value he placed on the term *often*: the greater that value, the less likely he ought to have been to check that item true. Thus a negative correlation would be expected between the SSI term *often* and score on the N item. Similarly, a composite of SSI terms should be expected to correlate negatively with the set of personality items (comprising all or part of a personality scale) that made use of those terms. Accordingly, a set of semantic style composites was constructed, and their correlations with the personality scales examined. (The rationale behind additional analyses, suggested by the data, will be discussed in a later section.)

### C. RESULTS

Three *Semantic Style* (SS) measures were formed by noting which SSI terms were used in a given personality scale or subscale, then constructing a weighted sum of those terms; the weighting was based on the number of times a particular term was employed by the personality scale. Table 1 lists, for the measures SS(MC Neg), SS(E), and SS(N), as well as for other measures to be described subsequently, the terms used and their respective weightings. Thus, SS(E) was the sum of the SSI score for *almost always*, twice the score for *usually*, and the score for *often* (three SSI terms were used a total of four times in the E scale). No SS(MC Pos) measure is shown in Table 1; the reason is that this composite, consisting almost exclusively of the extremely low variance terms *always* and *never*, would necessarily have very low variance itself and, therefore, could not have been expected to have correlates of any magnitude. [Actually, SS(MC Pos) was constructed, and the expected absence of significant correlations confirmed.] Separate SS measures were, of course, required for the MC subscales, since the subscales presumably reflect opposite poles of the social desirability variable; total MC is, in effect, computed by the subtraction of one subscale from the other. This balanced-keying issue was not met with E and N because nearly all the items for these two scales are keyed in the same, positive direction, and the very few that are not do not employ SSI terms.

TABLE 1  
SEMANTIC SCALING INVENTORY QUANTIFIER TERMS AND THE WEIGHTINGS USED  
IN CONSTRUCTION OF SEMANTIC STYLE VARIABLES

Quantifier terms	Mean	SD	Semantic style variables				
			SS (MC Neg)	SS(E)	SS(N)	SS (F Low)	SS (F High)
Always	99	1.1			1		1
Almost always	92	5.6		1			1
Regularly	81	10.2					1
Mostly	78	11.6					1
Frequently	74	9.6					1
Usually	74	11.9		2	1		1
Often	73	9.5		1	4		1
Generally	66	13.8					1
Fairly often	63	9.8					1
Sometimes	46	13.0	5		3	1	1
At times	42	11.9	4		2	1	1
Occasionally	40	13.2	4			1	1
Now and then	38	13.4			2	1	1
Seldom	18	7.9				1	
Hardly ever	14	7.4				1	
Rarely	13	6.1				1	
Almost never	7	5.0				1	
Never	1	1.1				1	

Note: Definitions of abbreviations: SS = semantic style; MC Neg = Marlowe-Crowne subscale of negative items; E = Extraversion scale; N = Neuroticism scale; F = factor.

Also listed in Table 1 are means and standard deviations for the SSI terms. SS(MC Neg) consisted of intermediate-level terms, SS(E) of relatively high ones, while the range of terms for SS(N) was quite broad. Distributions of all three composites were unimodal and approximately symmetric.

Correlations between the personality scales and subscales and the SS measures are given in Table 2. The relation between MC Neg and SS(MC Neg) was very significant; parallel associations for E and N were nonsignificant, though in the predicted direction. MC Pos and the full MC scale are included in the table, even though there are no corresponding SS measures. Notwithstanding this fact, the correlation between MC and SS(MC Neg) just missed significance at the .05 level, a consequence of the relation between MC(Neg) and the scaling variable (this coefficient was positive because MC Neg contributed negatively to the total MC).

Interestingly, there were two directionally appropriate, significant relations between personality scales and noncorresponding SS measures. That SS(N) should have predicted MC Neg was not startling, since the SS measures for MC Neg and N were highly related ( $r = .85$ ), partly because of shared SSI items. SS(E)'s prediction of N, however, was a surprising finding.

The preceding analysis had examined the relation between semantic style

TABLE 2  
CORRELATIONS BETWEEN SEMANTIC STYLE VARIABLES  
AND PERSONALITY MEASURES

Personality measures	Semantic style variables		
	SS (MC Neg)	SS(E)	SS(N)
Scales and subscales			
MC	.30	.18	.28
MC Pos	.09	.09	.08
MC Neg	-.44**	-.15	-.40**
E	-.11	-.12	-.11
N	-.12	-.32*	-.18
Subsets			
MC Neg'	-.48**	-.20	-.44**
N'	-.13	-.36*	-.21

Note: Italicized coefficients are those between corresponding variables, for which negative values were predicted. Definitions of abbreviations are as follows: MC = Marlowe-Crowne Social Desirability Scale; MC Pos = subscale of positive items; MC Neg = subscale of negative items; E = Extraversion scale; N = Neuroticism scale; SS = semantic style.

\*  $p < .05$  (two-tailed).

\*\*  $p < .01$ .

and existing standard measures of personality. The question next arose whether stronger evidence of semantic style-content association might be seen if just those personality items employing SSI terms were included in the content measures. Modified MC Neg and N scales thus were formed, subsets of the original measures, and these then correlated with the SS variables. The MC Neg' subset was quite similar to the overall MC Neg: 13 of the original 15 items used SSI terms and therefore were retained. The modified N' scale used 13 of the original 19 N items. A modified E' scale would have consisted of only four (dichotomous) items, with doubtless a substantial drop in reliability, and hence could not have been very informative.

Correlations between these two new content measures and the SS scales are shown in Table 2. As predicted, the coefficients did increase in magnitude, though not by any appreciable amount. Only a slight increase could have been expected for MC Neg', of course, because of its close similarity to MC Neg. In the case of N', however, the potential for improvement was greater.

A third analysis stemmed from the previous, unexpected finding that one semantic style measure, SS(E), correlated significantly with a noncorresponding personality scale, N. This suggested the possibility that a more-or-less general dimension of semantic style exists that relates to diverse personality measures, and that the true extent of content-style relation



might not be revealed by restricting an SS measure to terms explicitly used in the given content measure.

To aid in this inquiry, a principal components factor analysis was performed on the 20 SSI terms (the  $N$  here was 81, the present sample augmented by other, similar Ss who had completed the SSI). Although some indication of a general semantic style factor was given by the finding that all but two terms loaded in the same direction on the first principal component, several of these directionally consistent variables had low-value loadings. Furthermore, the second principal component was nearly as large as the first. The two together accounted for 37% of the total variance, and each was substantially greater than succeeding components. Consequently, the first two components were rotated, by the Varimax method.

One rotated component was characterized by terms with low means (*seldom, rarely, almost never*), the other by terms with relatively high means (*fairly often, often, usually*).<sup>4</sup> Four terms with means of intermediate level, *sometimes, at times, occasionally, and now and then* (and only these four), had substantial representation on both components.

Approximate measures of these two factors, SS(F Low) and SS(F High), were formed by the unweighted summing of SSI terms having factor loadings  $\geq .35$ . Table 1 shows the exact composition of these two new variables. That four intermediate-mean terms were shared by these two factor scores suggested another logical basis for construction of composite variables: the two factor-based measures *excepting* the four common terms, denoted SS(F Low)' and SS(F High)', and the sum of the four intermediate terms themselves, SS(F Medium)'. All five of these new SS variables, like the previous three, were unimodal and roughly symmetric in their distributions.

Table 3 contains the personality scale correlates of the new, factor-based SS variables. One sees that both the two broader measures, SS(F Low) and SS(F High), related significantly, with the latter having the more pervasive association. Correlations involving the other three narrower SS measures help clarify the content-style relations now seen from a number of perspectives. The medium-level SSI terms of SS(F Medium)' appear to have been instrumental in predicting MC Neg, whereas it was the high-level terms of SS (F High)' that seem to have been responsible for association with N. The low-mean terms of SS(F Low)' bore no reliable relation to the personality variables.

<sup>4</sup> It is not the case that this "level" basis for factor differentiation was artifactual, resulting from sets of oppositely skewed terms. With the exception of the low variance, extreme-mean quantifiers *always* and *never*, all terms had reasonably symmetric distributions.



TABLE 3  
CORRELATIONS BETWEEN PERSONALITY SCALES AND FACTOR  
ANALYSIS-BASED SEMANTIC STYLE VARIABLES

Personality measures	Semantic style variables				
	SS (F Low)	SS (F High)	SS (F Low)'	SS (F Medium)'	SS (F High)'
MC	.24	.31*	-.00	.30	.23
MC Pos	.00	.12	-.15	.08	.12
MC Neg	-.44**	-.37*	-.20	-.45**	-.19
E	-.19	-.07	-.27	-.09	-.03
N	-.04	-.32*	-.18	-.14	-.38*

Note: Definitions of abbreviations: MC = Marlowe-Crowne Social Desirability Scale; MC Pos = subscale of positive items; MC Neg = subscale of negative items; E = Extraversion scale; N = Neuroticism scale; SS = semantic style; F = factor.

\*  $p < .05$  (two-tailed).

\*\*  $p < .01$ .

Supplementary results of the study were as follows. A dichotomous (0, 1) sex variable was formed and correlated with all other variables. None of these point-biserial coefficients reached even the  $p < .10$  (two-tailed) level of .26, nor did the significant SS-content correlations of Tables 2-3 change appreciably when this sex variable was partialled out.

Variants of the original SS scales [SS(MC Neg), SS(E), and SS(N)] were formed by the unweighted summing of constituent items, and correlations between these nonexplicitly weighted SS measures and the personality variables were highly similar to those reported in Table 2. Coefficient alpha reliabilities were computed for these new SS measures and found to be, respectively, .76, .74, and .75. Reliability coefficients for the five factor-based measures, in the left-to-right order of Table 3, were .79, .80, .77, .85, and .69. (These last correlations may have been spuriously high, to a degree, because the 40 Ss of this study formed half the data base for the factor analyses.)

#### D. DISCUSSION

For two of the three personality scales (as well as for the MC Neg subscale) significant correlations in the predicted direction were found with one or more semantic style variables. Moreover, the true extent of SS-content association might have been underestimated, for the SSI was not constructed specifically for the task to which it was put in the present research. Although 13 of the 15 MC Neg items employed SSI modifiers, and 13 of the 19 N items, only four E items were so represented—and it was in fact only the E scale that showed no significant relationship with an SS measure. It is conceivable that, had the SSI included the quantifier terms

found in eight additional E items, a significant relationship would have emerged for E as well as for MC and N. Some (admittedly modest) support for this speculation is given by the slight increase in correlational values shown (Table 2) when the subsets MC Neg' and N' were used, rather than their counterparts MC Neg and N.

On the other hand, it seems not the case that the best SS predictor for a given content variable is one necessarily consisting of those modifier terms, perhaps weighted by frequency of usage, explicitly found in the content measure. This assertion certainly appears to hold for N, where the greatest correlation involved a factor analysis-based SS measure. In the case of MC Neg, there was advantage neither one way nor the other. (However, SS(MC Neg) and SS(F Medium)' were highly associated.)

With regard to factor analytic measures, it seems clear that a single semantic style dimension is unlikely to be the optimal index. Although the various SS composites were related almost uniformly positively with one another (25 of the 28 coefficients), and in some instances to very great degree, certain distinctions are nevertheless obvious. Speaking to this point are both the absence of a dominating general factor and the differential predictability of the level-based factor scores [the SS(F)' measures]. Patently, this specificity-generality issue needs further exploration.

A more important question than the degree to which the particular content measures here examined are related to semantic style is the extent to which the latter may permeate questionnaire or verbal measures generally. One might not be surprised to find the scope extensive, since the three personality scales of this study were not selected with semantic style considerations in mind. E and N were chosen for administration because they measure two quite pervasive, broad dimensions of personality. MC was included because it is a frequently employed measure of the important personality variable, social desirability. The three scales tap distinct, at least relatively independent constructs; the largest interscale correlation in the present study was that between E and N, an atypically high  $-.40$ .

The finding of appreciable semantic style variance in a given questionnaire ought not, of course, be a terminal research aim. One would wish to purify the content measure in some way so that the variable sought could be more adequately assessed. Among the courses imaginable are the following. First, items could be written (or rewritten) without the use of quantifier terms. This approach, however, is of doubtful promise. The point to inclusion of quantifiers in the first place is to reduce item ambiguity. Items with quantifiers may be ambiguous—a thesis of this paper—but items without quantifiers

are typically more ambiguous still. Related to this issue is the likelihood that Ss would furnish, implicitly, their own quantifier terms should such terms not actually be present. Here, then, added to the variation inherent in a particular quantifier would be variation stemming from individual differences in choice of that quantifier.

A more fruitful tack would seem to be the writing of questionnaire items<sup>5</sup> in such a way that all make use of certain selected quantifiers, and to accompany presentation of the resulting questionnaire with an SSI-like inventory assessing the relative value of these quantifiers. With semantic style "contamination" thus systematically identified, one could partial out (or otherwise isolate) the stylistic component to arrive at an SS-free measure of the content sought. As remarked previously, the present SSI might well not be the optimal instrument, and investigation of other kinds of psychometric devices would be a desired preliminary endeavor.

Another possibility would be use of modifiers in balanced scales (half the items keyed in one direction, half in the other), a tactic often suggested for the control of acquiescence style. This would appear in any given context only a partial solution and one whose applicability is not universal (*cf.* the SS-MC correlations above). For some purposes, it may be difficult or impossible to write adequate reversal items (witness the authoritarianism-acquiescence literature), and reversal items may not partake of the same degree of semantic style as the originals.

Still another approach would be to replace modifier terms with explicit quantitative expressions: e.g., "On average, I am moody *one* (*two, three*, etc.) days a week." Although this approach may well merit systematic exploration, it would seem applicable only to items employing frequency modifiers. (How, e.g., would one deal with an intensity format item such as, "My heart beats fast when giving a speech"?)

Previous mention of acquiescence style suggests another point of discussion: the relation of this response style variable to semantic style. Though alike in some respects, the two notions nonetheless are conceptually distinct. Specifically, one can imagine both factors influencing an S faced with the item, "My mood often goes up and down." The semantic style factor could furnish an increment of endorsement probability based on the meaning S ascribed to *often*, the acquiescence factor an increment based on his general disposition to respond affirmatively<sup>6</sup> to this (or any other) item, apart from

<sup>5</sup> Such items need not be of dichotomous form. Gradated response formats possess certain advantages (9).

<sup>6</sup> Bentler, Jackson, and Messick's (1) "agreement acquiescence"; the difference between semantic style and their "acceptance acquiescence" is still more pronounced.

the particular quantitative meaning given *often*. This conceptual distinction aside, the actual relation between the two variables remains, of course, an empirical matter.

A final issue for discussion is the degree to which, should semantic style be a reliable source of questionnaire variance, that variance is truly stylistic or content-irrelevant. This paper has presented semantic style as in fact a noncontent method variable, and the results of this study seem consistent with this stylistic interpretation. The SSI presents an utterly context-free task to the subject: namely, that he determine where, along a continuum marked only at the extremes, a diversity of quantifier terms should be placed. Neither of the constructs purported to be measured by MC and N has been described by investigators as having a semantic style-like aspect, yet both MC and N correlated significantly with an SS measure. Again, however, the issue is empirical, and more impressive support for the stylistic view will be given by the correlation with semantic style variables of a number of diverse-content measures.

#### REFERENCES

1. BENTLER, P. M., JACKSON, D. N., & MESSICK, S. Identification of content and style: A two-dimensional interpretation of acquiescence. *Psychol. Bull.*, 1971, **76**, 186-204.
2. CROWNE, D. P., & MARLOWE, D. A new scale of social desirability independent of psychopathology. *J. Consult. Psychol.*, 1960, **24**, 349-354.
3. EYSENCK, S. B. G., & EYSENCK, H. J. The measurement of psychoticism: A study of factor stability and reliability. *Brit. J. Soc. & Clin. Psychol.*, 1968, **7**, 286-294.
4. FISKE, D. W. *Measuring the Concepts of Personality*. Chicago: Aldine, 1971.
5. HAKEL, M. D. How often is often? *Amer. Psychol.*, 1968, **23**, 533-534.
6. MEEHL, P. E. The dynamics of "structured" personality tests. *J. Clin. Psychol.*, 1945, **1**, 296-303.
7. SHEPPARD, D. The adequacy of everyday quantitative expressions as measurements of qualities. *Brit. J. Psychol.*, 1954, **14**, 40-50.
8. SIMPSON, R. H. The specific meanings of certain terms indicating different degrees of frequency. *Quart. J. Speech*, 1944, **30**, 328-330.
9. STRAHAN, R. Subject satisfaction with questionnaire assessment: A function of response format. *Proc. 79th Ann. Conven. APA*, 1971, **6**, 125-126.

Department of Psychology  
Iowa State University  
Old Botany Hall  
Ames, Iowa 50010



# RELATIONS AMONG PHYSICAL ATTRACTIVENESS, BODY ATTITUDES, AND SELF-CONCEPT IN MALE AND FEMALE COLLEGE STUDENTS\*<sup>1</sup>

*Eastern Michigan University*

RICHARD M. LERNER,<sup>2</sup> STUART A. KARABENICK, AND JOYCE L. STUART

## SUMMARY

One hundred eighteen male and 190 female college students rated 24 body characteristics in terms of (a) how satisfied they were with each of these characteristics of their own body; (b) how important each part was in determining their own physical attractiveness; and (c) how important each part was in determining opposite-sex physical attractiveness. In addition, all subjects responded to a short self-concept scale. Results indicated that (a) males and females rated the importance of the body characteristics for both their own and opposite-sex physical attractiveness in a markedly similar manner; (b) mean satisfaction ratings were significantly related to self-concept among both males and females, and different body parts were differentially significantly related to self-concept; (c) weighting satisfaction ratings by corresponding importance ratings did not increase the satisfaction/self-concept relation; but (d) males' mean importance of opposite-sex body characteristics ratings were significantly related to self-concept, while the corresponding relation was not significant for females; and (e) subjects' physique type was not related to self-concept. The role of sex-linked body stereotypes and the differential role of the body in the personality development of males and females were discussed.

## A. INTRODUCTION

Recent research (e.g., 3, 5) has demonstrated a relation between physique type and self-concept; specifically, "chubby" male children and adolescents have unfavorable attitudes toward their physique, while those with "average" physiques show positive attitudes toward their bodies. These findings

\* Received in the Editorial Office on June 30, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> This research was supported in part by NSF Grant No. 25652-26.

<sup>2</sup> Reprint requests should be sent to the first author at the address shown at the end of this article.



suggest a more general relation between body attitudes and self-concept: Persons holding positive body attitudes should have generally more positive self-concepts than those with negative body attitudes.

A study by Rosen and Ross (7) provides some support for this hypothesis. These authors report that among 82 undergraduate subjects, mean "satisfaction-for-body-parts" scores were related to measures of self-concept. Furthermore, this relation was increased when the authors considered only satisfaction scores for those body parts that were rated by subjects as above their mean "importance-for-body-parts" ratings. However, this study did not differentiate between male and female subjects. This omission is crucial in view of the considerable evidence (e.g., 4) that the role of the body in personality development is different for girls than for boys. It may very well be the case that the relation between body attitudes and self-concept would hold for females but not for males. Alternatively, it might be predicted that the relation exists for both sexes, but that it is greater for females than for males.

Furthermore, it would be important to consider not only the overall relation between body attitudes and self-concept, but also the relative contribution of selected aspects of the body in determining self-concept. This consideration would involve an inspection of the relative importance ratings of selected body parts and the relation between satisfaction ratings of these selected parts and self-concept. Moreover, in view of possible sex differences, it would be crucial to consider the degree to which males and females agree and/or disagree about the relative importance of certain body parts in determining their own attractiveness. Accordingly, another aspect of this issue involves the importance that persons place in various body parts for determining the attractiveness of the opposite sex. It would be important to examine the relation of such ratings to opposite-sex importance, same-sex importance, and same-sex satisfaction of body characteristics to self-concept. In sum, the present study was designed to extend the findings of previous research on body attitudes and self-concept (e.g., 3, 5) to older groups of males and females, to determine the generalizability of the previous results of Rosen and Ross (7), and to obtain a more differentiated understanding of the relations among body attitudes, importance of physical attractiveness, and self-concept.

## B. METHOD

### 1. Subjects

Subjects consisted of 118 males and 190 females (total  $N = 308$ ) enrolled in various undergraduate psychology courses at Eastern Michigan

University in the Spring of 1972. The mean age of the males was 23.3 years ( $SD = 4.4$  years), and of the females was 21.9 years ( $SD = 5.1$  years). Ninety-five percent of the sample were white. The small representation of black subjects precluded any analyses involving race differences.

## 2. Procedure

Subjects were tested in groups within their classrooms. Each subject was given a test booklet, including a cover sheet asking for answers to various demographic questions (e.g., age, sex, height, weight), and this was followed by a series of randomly ordered scales. Responses to the scale questions were recorded by the subjects on answer sheets enclosed in the booklets.

## 3. Scales

Subjects responded to four scales, three of which (Numbers 1, 2, and 4) were adopted from Rosen and Ross (7).<sup>3</sup> Scales 1, 2, and 3 each presented a list of 24 body characteristics. With Scale 1 subjects rated each characteristic on a five-point Likert scale in terms of how "satisfied" they were with the appearance of each of these characteristics of their own body. Response alternatives ranged from "5" = "very satisfied" to "1" = "very dissatisfied." With Scale 2, subjects rated each body characteristic in terms of how "important" each characteristic was to them (i.e., how important each characteristic was in determining how physically attractive they themselves were). Response alternatives ranged from "5" = "very important" to "1" = "very unimportant." With Scale 3, subjects rated each body part in terms of how important each characteristic was to them in determining the physical attractiveness of members of the opposite sex. Response alternatives were identical to those of Scale 2. Scale 4, originally derived from Gough's (2) Adjective Check List, consisted of a short self-concept scale comprised of 16 bipolar dimensions (e.g., "mature-immature," "capable-not capable"). Scores on these basically evaluative items have been shown by Rosen and Ross to predict body attitude scale scores, and one objective of the present research was to ascertain relations of this scale to other scales within particular subject subgroups (e.g., males and females). Response alternatives for Scale 4 ranged from "5" = "word on left end of the continuum is most like me" to "1" = "word on right end of the continuum is most like me."

<sup>3</sup> Copies of all scales may be obtained from the first author at the address shown at the end of this article.

## C. RESULTS

1. *Importance of Body Characteristics for Own- and Opposite-Sex Physical Attractiveness*

Table 1 presents the mean own- and opposite-sex importance ratings for each of the 24 body characteristics, separately for male and female subjects. Data in the table have been coded so that "5" represents the very important and "1" represents the very unimportant ends of the scales (Scales 2 and 3, respectively). Also presented in Table 1 are the rank orders of the means.

First, the data indicate a strikingly close correspondence between all the mean ratings. The rank-order correlation for all body parts between the males' ratings of "own importance" and the females' ratings of these parts for both "own importance" and "opposite-sex importance" was  $+ .96$  and  $+ .94$ , respectively. Similarly, the males' ratings of these body parts for "opposite-sex importance" was correlated  $+ .94$  with females' "own importance" ratings, and  $+ .96$  with females' "opposite-sex importance" ratings of these body characteristics. Thus, both males and females rate the importance of the various body parts for both own and opposite-sex physical attractiveness in a markedly similar way.

The ranking similarity clearly suggests that most of the selected body characteristics play a similar role for both males and females in determining their judgment of their own physical attractiveness and the physical attractiveness of opposite-sex members. For example, as seen in Table 1, "general appearance" is ranked first in all sets of ratings, and the mean ratings for this characteristic are not significantly different among the four rating scales. Other characteristics given similar rankings are "teeth," "ears," "ankles," "neck," and "body build."

Perhaps the most interesting consistency seen in Table 1 is the relation between a sex's own-importance-of-body-characteristics rankings and those of the opposite sex. Both males and females agree that what they deem important for determining their own attractiveness is in agreement with what members of the opposite sex deem important in finding them physically attractive.

Despite the marked similarities, however, there are some notable differences in the ratings and rankings of body characteristics. Males consider the "shape of legs," "hips," and "thighs" more important in determining opposite-sex physical attractiveness than in determining their own physical attractiveness. Females appear to agree with these judgments; they rate

TABLE 1  
MEANS AND RANKS OF IMPORTANCE OF SELECTED BODY CHARACTERISTICS FOR OWN  
AND OPPOSITE-SEX PHYSICAL ATTRACTIVENESS FOR MALES (N = 118) AND  
FEMALES (N = 190)

Body characteristics	Males' importance						Females' importance					
	Own			Opposite-sex			Own			Opposite-sex		
	$\bar{X}$	SD	Rank	$\bar{X}$	SD	Rank	$\bar{X}$	SD	Rank	$\bar{X}$	SD	Rank
Facial complexion	4.2	.7	3.5	4.3	.8	3.0	4.1	.9	4.5	3.9	.9	4.0
Ears	2.5	1.1	23.0	2.5	1.0	23.0	1.9	1.0	24.0	2.0	1.1	23.0
Chest	3.6	.8	11.0	3.9	.9	9.0	3.7	1.1	9.0	3.1	1.1	10.5
Profile	3.8	.8	7.5	3.9	.8	9.0	3.3	1.1	14.0	3.0	1.1	13.0
Weight												
distribution	4.2	.8	3.5	4.2	.8	4.0	4.3	.7	3.0	4.0	.8	2.5
Eyes	3.8	1.1	7.5	3.7	1.2	13.0	3.7	1.1	9.0	3.6	1.1	7.5
Height	3.5	1.1	14.0	3.3	1.1	17.5	3.3	1.1	14.0	3.6	1.1	7.5
Ankles	2.0	1.1	24.0	2.3	1.2	24.0	2.1	1.2	23.0	1.6	.9	24.0
Waist	3.7	.9	9.0	3.9	.9	9.0	3.6	.9	11.0	3.0	1.0	13.0
Arms	3.4	1.0	16.0	3.1	.9	20.5	2.8	1.0	19.0	2.9	1.0	15.0
Shape of legs	3.2	1.1	19.0	3.8	1.1	12.0	3.8	1.0	6.5	2.6	1.2	19.0
General appearance	4.6	.7	1.0	4.5	.8	1.0	4.7	.6	1.0	4.5	.8	1.0
Hips	3.3	1.0	17.5	3.9	.9	9.0	3.7	.9	9.0	2.8	1.2	16.5
Shoulder width	3.6	.9	11.0	3.1	1.1	20.5	2.6	1.1	21.5	3.1	1.1	10.5
Mouth	3.5	1.0	14.0	3.6	1.0	15.0	3.3	1.0	14.0	3.3	1.0	9.0
Neck	3.1	1.1	20.5	3.2	1.1	19.0	2.6	1.1	21.5	2.5	1.1	21.0
Teeth	4.1	.9	5.5	4.0	.9	6.0	3.8	1.0	6.5	3.7	1.1	6.0
Nose	3.6	.9	11.0	3.6	.9	15.0	3.2	1.0	16.5	3.0	1.1	13.0
Chin	3.3	.9	17.5	3.3	1.0	17.5	2.7	1.1	20.0	2.6	1.0	19.0
Hair texture	3.5	1.1	14.0	3.6	1.0	15.0	3.2	1.3	16.5	2.8	1.2	16.5
Body build	4.1	.8	5.5	4.1	.8	5.0	4.1	.9	4.5	3.8	1.0	5.0
Hair color	3.0	1.2	22.0	3.0	1.2	22.0	2.9	1.3	18.0	2.2	1.2	22.0
Thighs	3.1	1.1	20.5	3.9	1.0	9.0	3.4	1.0	12.0	2.6	1.1	19.0
Face	4.3	.7	2.0	4.4	.9	2.0	4.4	.6	2.0	4.0	.9	2.5



and rank these body parts as more important in judging their own physical attractiveness than in evaluating the physical attractiveness of males. Conversely, females consider "height" and "width of shoulders" more important for judging the physical attractiveness of males' bodies than their own bodies. Consistent with these judgments, males rate these body characteristics more important in evaluating their own attractiveness than the attractiveness of females.

## 2. *Satisfaction with Body Characteristics and Self-Concept*

To test the hypothesis that a general relation exists between body attitudes and self-concept, the mean male and female ratings of their own body characteristics satisfaction scores (Scale 1) were correlated with their self-concept scores (Scale 4). As shown in Table 2, significant correlations ob-

TABLE 2  
CORRELATIONS BETWEEN IMPORTANCE, SATISFACTION, AND SELF-CONCEPT RATINGS  
AMONG MALES (BELOW DIAGONAL) AND FEMALES (ABOVE DIAGONAL)

Variables	1	2	3	4	$\bar{X}$	SD
1. Mean own-importance						
2. Mean opposite-sex importance	.58**	.61**	.13	.08	3.5	.6
3. Mean satisfaction	.21*	.14	.05	.11	3.6	.6
4. Self-concept	.09	.25**	.33**	.43**	4.0	.6
$\bar{X}$	3.4	3.1	3.9	44.1	45.2	7.7
SD	.6	.6	.5	8.3		

\*  $p < .05$ .

\*\*  $p < .01$ .

tained with both the males ( $r = .33$ ,  $df = 116$ ,  $p < .01$ ) and the females ( $r = .43$ ,  $df = 188$ ,  $p < .01$ ). This finding supports the hypothesis that the degree of positive self-concept increases with the degree of satisfaction (i.e., positive attitude toward) one's body characteristics. The relationship is present for both males and females; furthermore, although the relation is apparently more pronounced for females than for males, from Table 2 it can be seen that the difference between these correlations is not due to either mean rating differences or differences in the variability of the ratings between males and females. Nevertheless, the difference between the correlation coefficients is not significant ( $z = .99$ ,  $df = 306$ ,  $p < .15$ ).

## 3. *Relationship Between Satisfaction with Selected Body Characteristics and Self-Concept*

In addition to the general relation between mean body characteristics satisfaction and self-concept, the results of previous research (7) have sug-



gested that satisfaction with specific body parts may be differentially related to self-concept. To test this notion in the present study, the satisfaction scores for each body characteristic were separately correlated with self-concept, for males and females respectively. The results of this analysis are presented in Table 3. With males, 10 characteristics were highly correlated with self-concept ( $p < .01$ ), seven items were moderately correlated ( $p < .05$ ), and seven characteristics were not correlated ( $p > .05$ ). Alternatively, with females, although more characteristics (16 characteristics) were highly

TABLE 3  
RELATIONS BETWEEN SATISFACTION RATINGS FOR EACH BODY CHARACTERISTIC AND  
SELF-CONCEPT FOR MALES AND FEMALES

Category	Males ( $N = 118$ )		Females ( $N = 190$ )	
	Body characteristic	$r$	Body characteristic	$r$
Highly correlated ( $p < .01$ )	Facial complexion	.36	Facial complexion	.21
	Distribution of weight	.26	Distribution of weight	.26
	Waist	.25	Waist	.23
	Nose	.28	Nose	.18
	Body build	.26	Body build	.42
	Face	.27	Face	.26
	Thighs	.24	Thighs	.28
	Shape of legs	.27	Chest	.25
	Teeth	.25	Profile	.30
	Hair Texture	.30	General appearance	.39
			Eyes	.23
			Height	.21
			Ankles	.27
			Hips	.30
			Chin	.21
			Hair-color	.22
Moderately correlated ( $p < .05$ )	Width of shoulders	.19	Width of shoulders	.16
	Profile	.21	Ears	.16
	General appearance	.22	Shape of legs	.17
	Eyes	.20		
	Mouth	.21		
	Neck	.18		
	Hair color	.20		
Not correlated (ns)	Ears	.04	Arms	.11
	Chest	.15	Mouth	.11
	Height	.00	Neck	.13
	Ankles	.04	Teeth	.10
	Arms	.11	Hair texture	.04
	Hips	.16		
	Chin	.09		

correlated with self-concept, fewer were moderately correlated (three) and not correlated (five). Moreover, for both sexes the mean importance ratings of all characteristics in the highly correlated category was higher than the mean importance ratings of the characteristics in the other two categories, respectively. Thus, for both sexes satisfaction with different body parts is differentially related to self-concept. In addition, consistent with the above results, 68% of the 19 body parts which were significantly related to self-concept among females were also so related among males, while 76% of the 17 body parts which were significantly related to self-concept among males were also significantly related to self-concept among females.

#### 4. "Importance" and "Satisfaction" as Related to Self-Concept

Rosen and Ross (7) reported that the relation between satisfaction with body characteristics and self-concept was increased when satisfaction ratings were considered in the context of the "importance" ratings of those body parts. To test this relation in the present study, each subject's satisfaction rating for each body part was linearly weighted by his corresponding rating of the importance of that body characteristic for judging "own physical attractiveness." The means of these weighted satisfaction scores were then correlated with the person's self-concept score. This was done separately for males and females. The correlation coefficients between satisfaction and self-concept did not increase when the weightings were used (males  $r = .33$ ; females  $r = .44$ ). It should be noted that the failure of the weighted satisfaction ratings to increase the correlations over the unweighted ratings is not due to a high relationship between satisfaction and importance scores. As shown in Table 2, low correlations between mean satisfaction and mean own-importance scores for both females ( $r = .13$ ,  $df = 188$ ,  $p < .10$ ) and males ( $r = .21$ ,  $df = 116$ ,  $p < .05$ ) suggest that the importance scores represent a relatively separate source of variation that could have potentially contributed to the satisfaction/self-concept relation. Thus, the Rosen and Ross finding is not supported by the results of the present study.

Another aspect of the relation between importance and self-concept is that between self-concept and the importance of *opposite-sex* body characteristics. As shown in Table 2, the mean importance of opposite-sex body characteristics ratings was significantly related to self-concept for males ( $r = .25$ ,  $p < .01$ ), but not for females ( $r = .11$ , ns). It should also be noted that the mean opposite-sex importance ratings were significantly greater for males (mean = 3.6) than for females (mean = 3.1;  $t = 7.1$ ;  $df = 306$ ,  $p < .001$ ).

### 5. *Physique Type and Self-Concept*

The previous research (e.g., 3, 5) that established that average physiques are related to positive self-concepts, and that chubby physiques are related to negative self-concepts, indexed body type through visual ratings of independent observers. That is, physique type groups were comprised of those males who were rated as either "chubby" or "average" by two independent raters. Physique groups in the present study were formed, however, through the use of traditional anthropometrical indices.

First, the Ponderal Index (PI) of each male and female was determined by means of standard procedure (1); this score was then correlated with mean self-concept scores. Second, standard insurance company tables of normative weight ranges for males and females of different heights and ages (6) were used to form Fat, Average, and Thin physique groups. The process used followed the procedure employed by Schachter and his colleagues (e.g., 8). Accordingly, Fat males and females were those subjects whose weight was 15% greater than appropriate height and age norms; Thin males and females were those subjects whose weight was 15% less than appropriate height and age norms; Average build males and females were those who were not more than 15% above, or 15% below, appropriate age and height norms. The male and female Fat, Average, and Thin groups were assigned the ranks "1," "2," and "3," respectively, and these rankings were correlated with the mean self-concept scores for the males and females, respectively.

The correlations between either physique type index and self-concept were not functionally different from zero. The correlations for males between PI and self-concept, and for the "1," "2," and "3" rankings and self-concept, were .02 and .08, respectively. For females, these correlations were .03 and .09, respectively. Consistent with the conclusions of Domey *et al.* (1), the present results suggest that traditional anthropometric indices of physique seem to have little predictive validity for psychological variables.

### D. DISCUSSION

The results show that college-aged males and females maintain markedly consistent attitudes about the importance of various body characteristics in determining both their own physical attractiveness and that of the opposite sex. That is, characteristics as global as "general appearance," or as specific as "ankles," are judged by both sexes in the same way, regardless of whether the sex group is rating the importance of that character-

istic for their own or opposite-sex attractiveness. Moreover, for the few characteristics with which own importance ratings are discrepant from opposite-sex ratings, the differences are in accord with obvious sex-linked body stereotypes. For example, both males and females agree "width of shoulders" and "height" are more important for males than for females, while "hips," "shape of legs," and "thighs" are more important for females than for males.

The results also indicated that males' mean importance-of-body-parts ratings for female body characteristics were positively related to self-concept in males. The corresponding relation was not significant for females. These findings suggest that an integral proportion of the variation in males' self-concepts is accounted for by males' overall level of being attracted to the opposite sex. That is, part of the male role may be comprised of being attracted to females' bodies. This notion is expressed in stereotyped societal norms for maleness, as represented, for example, in the cliché that "a woman is as old as she looks, while a man is old when he stops looking." The norm for females, however, is *not* to be primarily concerned with the attractiveness of a man's body, but perhaps rather with "romance" (9). The significantly lower mean opposite-sex importance-of-body-characteristics rating for females, as compared with males, may be viewed as consistent with this interpretation.

Thus, a general implication of these findings is that studies which do not consider sex differences in the role of the body in personality development (e.g., 7) are making crucial omissions. General relations between body attitudes and self-concept, such as those reported by Rosen and Ross (7), will remain imprecise unless apparently ubiquitous and critical sex differences are considered. For instance, Rosen and Ross reported that the relation between body satisfaction and self-concept was increased when only body parts that subjects judged as highly important were considered. In the present study, however, when the variation contributed to this relation by both male and female scores was considered, no support for the presence of this relation was found.

Finally, in the present study there were no significant relations between either of two anthropometric indices of body type and self-concept. Such indices have been criticized for lack of reliability, validity, and psychological relevance (e.g., 1). Alternatively, indices of physique type which rely on simple visual inspection by two independently agreeing raters have perfect reliability and predictive validity, and have been useful in establishing relations between physique and personality (5). While anthropometric in-

dices of physique have the aura of mathematical sophistication about them, they seem to have little psychological utility; here visual indices of physique are more useful. After all, it may be irrelevant if someone is "anthropometrically" of average build if, in fact, he appears to be fat to others and to himself. It will be this "apparent" body type which will show significant relations to the personality and behavior development of the person (e.g., 3, 4, 5).

## REFERENCES

1. DOMEY, R. G., DUCKWORTH, J. E., & MORANDI, A. J. Taxonomies and correlates of physique. *Psychol. Bull.*, 1964, **62**, 411-426.
2. GOUGE, H. The Adjective Check List. Palo Alto, Calif.: Consult. Psychol. Press, 1952.
3. LERNER, R. M. "Richness" analyses of body build stereotype development. *Devel. Psychol.*, 1972, **7**, 219.
4. LERNER, R. M., & GELLERT, E. Body build identification, preference, and aversion in children. *Devel. Psychol.*, 1969, **1**, 456-462.
5. LERNER, R. M., & KORN, S. J. The development of body build stereotypes in males. *Child Devel.*, 1972, **43**, 912-920.
6. METROPOLITAN LIFE INSURANCE COMPANY. New weight standards for men and women. *Stat. Bull.*, 1959, **40**, 1-7.
7. ROSEN, G. M., & ROSS, A. O. Relationship of body image to self-concept. *J. Consult. & Clin. Psychol.*, 1968, **32**, 100.
8. SCHACHTER, S., GOLDMAN, R., & GORDON, A. Effects of fear, food deprivation, and obesity on eating. *J. Personal. & Soc. Psychol.*, 1968, **10**, 91-97.
9. STATON, T. F. Sex education for adolescents. In J. F. Adams (Ed.), *Understanding Adolescence*. Boston: Allyn & Bacon, 1968.

*Psychology Department*  
*Eastern Michigan University*  
*Ypsilanti, Michigan 48197*



The first part of the paper is devoted to the study of the properties of the function  $f(x)$  defined by the equation  $f(x) = \sum_{n=0}^{\infty} \frac{f_n(x)}{n!}$ , where  $f_n(x)$  is a function of the  $n$ -th order of the differential equation  $y^{(n)} = P_{n-1}(x)y^{(n-1)} + \dots + P_0(x)y$ . The function  $f(x)$  is shown to be a solution of the differential equation  $y^{(n)} = P_{n-1}(x)y^{(n-1)} + \dots + P_0(x)y$ .

$$f(x) = \sum_{n=0}^{\infty} \frac{f_n(x)}{n!}$$

$$f(x) = \sum_{n=0}^{\infty} \frac{f_n(x)}{n!}$$

## IN DEFENSE OF MEASURING ATTITUDES\*

*Camden College of Arts and Sciences, Rutgers University*

---

MARION STEININGER

---

### SUMMARY

To investigate the validity of personality measures, a 43-item liberalism-conservatism scale was administered to 86 college students and 40 noncollege adults. In both samples, scores were related to presidential preferences, fundamentalism, and church attendance; among the students, scores were also related to the wearing of jeans. It was not possible to determine why some items were better than others. It was argued that the data cannot be understood without a concept like "liberalism-conservatism," and that the scale turned out to be valid because it sampled widely from the domain of the construct, and because respondents did not fear the consequences of their answers, either in terms of self-image or in terms of the possible subsequent actions of someone in power.

---

### A. INTRODUCTION

The psychological literature currently includes considerable discussion of a recently revived controversy: how "general" are personality traits, and should one theorize in terms of them (1, 2, 4, 8)? This is considered to be in part an empirical question: given a measure of a trait, how well can one predict across different criteria for different groups of people? It is also a theoretical question: given some data, how should they be interpreted?

It is the thesis of this paper that at least in some cases one can predict well on the basis of attitude scale scores, and that personality concepts are necessary for an understanding of the data. This thesis is explored with reference to "liberalism-conservatism."

### B. METHOD

#### 1. Procedure

An "Opinion Survey" was constructed which contained 43 items conceptually or "intuitively" (4) related to social-political liberalism-con-

---

\* Received in the Editorial Office on June 30, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

servatism (LC). These items dealt with such diverse topics as capital punishment, tax laws, abortion, mental illness, sex morals, welfare, men's long hair, foreign policy, etc. To minimize response set, 23 of these items were liberally oriented, 20 conservatively oriented. An example of the former is, "We should pull out of Viet Nam immediately"; of the latter, "Capital punishment (electric chair, gas chamber) is necessary to decrease crime in this country." Respondents could "agree a lot," "agree a little," "disagree a little," or "disagree a lot." Their responses were scored  $-2$ ,  $-1$ ,  $+1$ , or  $+2$ , with  $+2$  always representing the most *conservative* response. A mean LC score was calculated for every respondent.

The Survey also included several criterion items. Respondents were asked whether they would prefer Nixon or McGovern; whether they would instead prefer Wallace; how often they wore jeans; how often they attended church or synagogue; and whether ". . . you believe the Bible is true exactly as written." Responses to the question about church attendance were scored as follows: "never," 0; "once or twice a year," 1; "a few times a year," 2; "once or twice a month," 3; "once a week," 4; "two to six times per week," 5. No one checked "every day." Responses to the question about wearing jeans were scored as follows: "almost all the time," 4; "much of the time," 3; "some of the time," 2; "rarely," 1. No one checked "never."

## 2. Subjects

Two groups responded to the Survey. The first sample consisted of 86 students taking introductory psychology. Participation was anonymous; it was not a course requirement, but no one refused to answer. The second sample consisted of 40 noncollege adults chosen on the basis of their presidential preference (20 for Nixon, 20 for McGovern). Each adult was known by the student who administered the Survey; otherwise, the adults were also guaranteed anonymity. Both groups' LC scores were divided into approximate thirds: the most liberal, the most conservative, and those in the middle.

## C. RESULTS

Table 1 shows the responses of the two samples to the criterion questions as a function of liberalism-conservatism scores.

In both samples, and for all criterion questions, the more conservative the LC scores, the more conservative the criterion responses. Thus, the more conservative the respondent, the more likely that (s)he favored Nixon;

TABLE 1  
PRESIDENTIAL PREFERENCE, BELIEF IN THE BIBLE, CHURCH ATTENDANCE, AND  
WEARING OF JEANS AS A FUNCTION OF LIBERALISM-CONSERVATISM (LC) IN  
A COLLEGE AND NONCOLLEGE SAMPLE

Criteria	College Sample			Noncollege Sample		
	Most liberal (N = 26)	Middle (N = 32)	Most conserv- ative (N = 28)	Most liberal (N = 11)	Middle (N = 17)	Most conserv- ative (N = 12)
Prefer Nixon	12%	38%	64%	9%	59%	75%
Prefer Wallace	4%	9%	32%	0%	0%	33%
Bible true as written	13%	13%	42%	9%	24%	50%
Church attendance	2.12	2.25	2.68	1.27	2.24	2.92
"Never"	24%	25%	7%	55%	18%	0%
Wear jeans	3.04	2.84	2.70	(not asked)		

Note: LC scores for each group were as follows: College sample—Most liberal, —1.41 to —.87; Middle, —.86 to —.47; Most conservative, —.46 to +.70. Noncollege sample—Most liberal, —1.67 to —1.00; Middle, —.99 to —.01; Most conservative, .00 to +1.46.

would have preferred Wallace; believed that the Bible is true as written; attended church, and more often; and wore jeans less often (in the college group).

Five item analyses were performed on the LC scale. For the college sample, the first three criteria in Table 1 were used to divide the sample (e.g., into Nixon and McGovern supporters); because of the small number of Wallace supporters in the noncollege sample, only the first and third criteria were used with them. The mean responses of the subgroups (e.g., of Nixon and McGovern supporters) were then calculated for each of the 43 items. The number of items that discriminated as expected (e.g., Nixon supporters having a more conservative mean response than McGovern supporters) was then counted separately in each item analysis. Of the 43 items, between 30 and 39 discriminated in the expected direction in the five item analyses. Since by chance, only  $21\frac{1}{2}$  would have done so in each analysis, the five chi squares ranged from 5.953 to 26.884 ( $df = 1$ ;  $p < .01$ , one-tailed).

Items that discriminated in the expected direction by a mean of more than .500 or in all five item analyses were compared with those that did not discriminate as well. Three possible differences between better and poorer items were considered: liberally *vs.* conservatively oriented; dealing with a specific issue *vs.* philosophical; and dealing with a current issue *vs.* traditional. No consistent differences between better and poorer items were found.

## D. DISCUSSION

The data of this study support the thesis that there are attitude constructs which can be measured, and that such measurement permits predictions for different samples and criteria. Without a construct like "liberalism-conservatism," furthermore, one cannot understand the data of this study or similar ones (7). It seems necessary, therefore, to examine the criticisms of such constructs with reference to the present study.

A basic argument against theorizing about personality is based on the low correlations between "behaviors" in different situations. The conclusion does not necessarily follow from the data, however [see, for example, Alker (1)]. Behaviors in all situations, including whichever two were correlated, are always complexly determined. Thus, a preference for Nixon is *not* related only to liberalism-conservatism, nor are the answers to the items on the questionnaire. Therefore, one can only try to answer the following question: How suitable were the selected situations and behaviors for eliciting the hypothesized characteristic? Not all situations and behaviors are suitable for eliciting "intelligence," for example; some conditions of observation are more "diagnostic" of intelligence than others, in which it is overshadowed by irrelevant (in that context) variables (6).

The same considerations apply to the area of attitudes. In the present study, the liberalism-conservatism scale was administered anonymously. Suppose that it had also been administered to the same sample with names revealed, and that the two sets of questionnaires could be matched. If the correlations between the two sets of scores were lower than the test-retest reliability of the scale, would one be forced to conclude that there is no trait of "liberalism-conservatism"? Not necessarily. In this case, answering anonymously is probably better diagnostically than answering with one's identity revealed. In this study, as in all others, neither the criteria (even behavioral ones) nor the items can perfectly diagnose the construct (3). It does not follow from this that it is futile to try to conceptualize and measure individual differences. One could argue instead that there should be greater effort to try to determine the most valid conditions of observation for the current conception of a trait. In the case of "liberalism-conservatism," it is likely that answering without concern for consequences (i.e., anonymously) permits the respondents to answer honestly; perhaps non-anonymously, under conditions of Lovell's "client contract" (5), the scale would also remain valid.

There is another characteristic of this scale which probably contributes to its validity: the respondents' answers to items seem at least somewhat



desirable to them, be they "liberals," "moderates," or "conservatives." In this the scale differs from similar measures of "adjustment" or "paranoia." While "liberalism" is socially desirable for many college students, this isn't true for all students, nor is it true for all items; furthermore, the social desirability of "liberalism" may not affect the rank order of students' mean scores, but only their absolute value. This is why each sample was divided separately into thirds.

The scale used in the present study does present some real challenges, however, which reveal clearly how minimal our understanding of our constructs is. The items for the scale were chosen on theoretical or "intuitive" (4) bases; nonetheless, it is not clear on any basis why some items were better than others. For example, items about wiretapping and capital punishment were among the better ones, while those about owning a gun and pornography were among the poorer ones. This is like finding that one vocabulary word "works" as part of an intelligence test while another doesn't. Usually, such a finding cannot be explained, which reveals that the relationship between "vocabulary" and "intelligence" is not understood yet. When no sense can be made of the different predicative powers of different items, this is very telling, since it reveals that we understand neither the construct nor the measuring instrument as well as we need to.

The greatest challenge facing theorists is not to obtain high correlation coefficients; it is to conceptualize all the coefficients obtained, be they high or low. Such conceptualization cannot proceed without consideration of the manner in which people organize their experiences. Ideas about capital punishment are related to ideas about other social-political issues. A good measuring instrument is one that enables or permits the individual to reveal the organization that he imposes on his experiences; it reveals something about the operation of the trait, be it intelligence or "liberalism-conservatism," even though performance on it is not an expression only of that trait. In the present instance, the scale was probably successful with different samples and different criteria because it sampled widely from the domain of the construct being investigated, and because respondents could answer relatively honestly, since they did not fear the consequences of their answers, either in terms of self-image or in terms of the possible subsequent actions of someone in power.

#### REFERENCES

1. ALKER, H. A. Is personality situationally specific or intra-psychically consistent? *J. Personol.*, 1972, 40, 1-16.

2. BEM, D. J. Constructing cross-situational consistencies in behavior: Some thoughts on Alker's critique of Mischel. *J. Personal.*, 1972, **40**, 17-26.
3. CRONBACH, L. J., & MEEHL, P. E. Construct validity in psychological tests. *Psychol. Bull.*, 1955, **52**, 281-302.
4. GOLDBERG, L. R. Some recent trends in personality assessment. *J. Personal. Assess.*, 1972, **36**, 547-560.
5. LOVELL, V. R. The human use of personality tests: A dissenting view. *Amer. Psychol.*, 1967, **22**, 383-393.
6. SMEDSLUND, J. Psychological diagnostics. *Psychol. Bull.*, 1969, **71**, 237-248.
7. STEININGER, M., & CAPALDI, M. On the appeal of Rizzo and Longstreth. *Psychol. Rep.*, 1972, **31**, 141-142.
8. WICKER, A. W. An examination of the "other variables" explanation of attitude-behavior inconsistency. *J. Personal. & Soc. Psychol.*, 1971, **19**, 18-30.

*Department of Psychology*  
*College of Arts and Sciences*  
*Rutgers University*  
*Camden, New Jersey 08102*

## A DEVICE FOR INVESTIGATING ADAPTATION TO SENSORY REARRANGEMENT\*

*Department of Psychology, Brandeis University*

---

JAMES R. LACKNER

---

### SUMMARY

A device is described that is a useful tool for studying the conditions under which adaptation to visual rearrangement occurs. This device has provisions for creating independent discordances among the visual, proprioceptive, and auditory directions of external targets. It also has provisions for automatically recording pointing responses to these targets during pre- and postexposure measurement conditions.

---

The visual-rearrangement experiment has in recent years become an important paradigm in the study of sensory-motor coordination. It is possible to generalize the notion of sensory rearrangement by applying it to other modalities in addition to vision. The advantage of such an approach is that it permits the systematic exploration and determination of those sensory and motor factors that are necessary before adaptive changes in sensory-motor coordination will occur in response to sensory rearrangement. The device to be described permits the creation of any desired discordance among the visual, auditory, and proprioceptive cues signalling the direction of an external target. An important feature of this device is that the movements made by the subject during exposure to sensory rearrangement are of the same type and in the same plane as those made during the pre- and postexposure measurement trials. This equality of response form between different conditions insures that the maximum adaptation occurring will be measured, since several studies have revealed that if the responses made in exposure and measurement trials involve different types of movements (continuous *vs.* ballistic), or different limbs (upper *vs.* lower) instead of the same, then less adaptation is demonstrated (1, 3, 5).

Figure 1 is a schematic illustration of the apparatus. It has provisions for introducing sensory discrepancies in apparent target location and provisions for recording pointing responses to these targets. The subject

---

\* Received in the Editorial Office on July 6, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

sits at one end of the device with his head stabilized by a bite board and looks through a pair of binocular goggles. The goggles have a form fitting mold that limits the subject's field of view to the aperture of the goggle lenses. These lenses are rotary (Risley) prisms that permit leftward or rightward lateral visual displacements of any desired value between 0 and 30 diopters.

From a level just below the subject's chin a shelf extends forward to the opposite end of the apparatus. Flush with the surface of this shelf at the far end of the apparatus is an  $80^\circ$  horizontal metal arc whose center of curvature corresponds to the center of the positioned subject's interocular axis. Holes have been drilled at  $1^\circ$  intervals in this arc, and metal pins that serve as visual targets for the subject can be inserted at any desired locations within  $30^\circ$  to the left or right of the subject's midline. When auditory presentations are desired as well, small speakers (Grason-Stadler ear-phones TDH-39Z) mounted on carrier-tubes can be placed, facing the subject's midline, on the top of the metal pins.

Another horizontal metal arc is located 8" beneath the arc just described. It has holes at  $1^\circ$  intervals that are in exact vertical alignment with the holes in the upper arc. The holes in the top arc are deep enough to hold firmly the  $\frac{1}{4}$ "-diameter metal pins but do not extend through its entire thickness; by contrast, the holes in the lower arc extend through its thickness. Beneath the lower arc is a horizontal plate that can be raised or lowered. When this plate is in its up position, its top surface is flush with the bottom of the arc immediately above it. Metal pins can then be placed into the holes in the arc, since the plate prevents them from dropping through. These pins can be placed in a vertical line with the target pins in the top arc or at any desired angular eccentricities.

For exposure conditions in which a visual-proprioceptive discordance is desired, the upper and lower pins are placed in alignment, and the rotary prisms in the goggles are adjusted to induce a displacement of the visual image of the upper pins. The subject reaches under the shelf and moves his hand below the apparent position of the top pins until it makes contact with the lower pins; in so doing, he receives veridical proprioceptive feedback about the true locus of the visual targets. After completion of the exposure trials, the bottom plate is lowered and the "feedback pins" drop down until their tips are below the top surface of the lower arc. As the bottom plate is lowered, an 8" high curved surface attached to the same shaft descends to the level previously occupied by the feedback pins. When

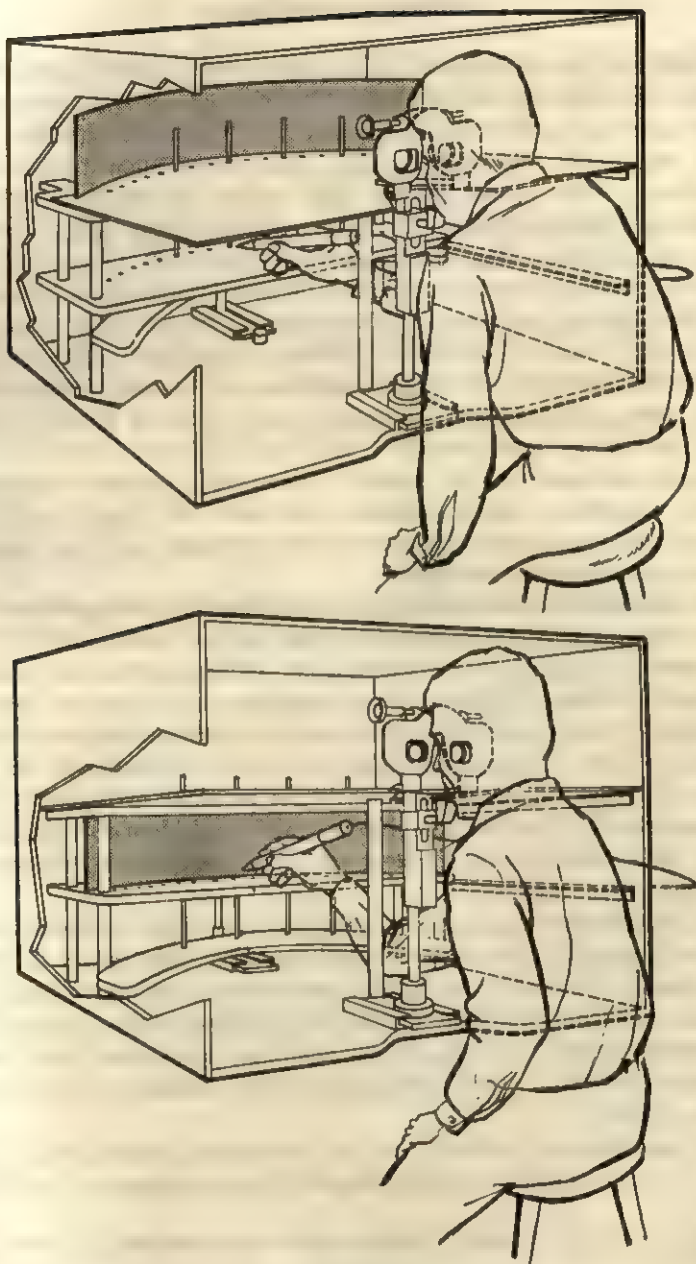


FIGURE 1

SCHEMATIC ILLUSTRATION OF THE EXPERIMENTAL APPARATUS

The upper portion of the figure depicts the exposure condition: the subject while looking at the upper pegs touches the lower pegs with the stylus. The lower portion of the figure depicts the pre- and postexposure measurement conditions: the subject while looking at the targets points just below their apparent position.



this surface is moved forward 1", its concave side is at exactly the same distance from the subject's midline as the visual targets in the upper arc.

The sector is made of lucite and has a 24" radius of curvature, an 8" height, and a  $\frac{1}{2}$ " thickness; its exact shape is maintained by a tension-strut system. The surface of the sector facing the subject is lined with conducting teledeltos paper. A voltage gradient exists from left-to-right across the teledeltos paper. Consequently, the value of the voltage at any horizontal position serves as a unique reference for that point.<sup>1</sup> In the pre- and postexposure measurement trials when the subject points to a visual (or auditory) target, he does so holding a pen-shaped stylus in such a fashion that the stylus tip is aligned with the end of his index finger. The tip of the stylus is spring-loaded and upon contact with the teledeltos paper is forced backwards, approximately  $\frac{1}{16}$ ", into the barrel of the stylus. The backward movement of the stylus tip depresses a miniature microswitch, thereby triggering a digital voltmeter that reads the voltage at the tip of the stylus. A digital printer connected to the voltmeter provides a permanent record of the subject's responses.

It is important to have each subject positioned in the apparatus in such a way that the midpoint of the line connecting his ocular orbits is always the center of curvature of the stimulus arc. Otherwise, it is not possible to standardize voltage readings at the recording arc in terms of degrees of visual angle. Precise placement of each subject is achieved by the appropriate positioning of his biteboard, up-down and in-out, in relation to the goggle frame. The height of the subject's chair can be raised or lowered to adjust the seated height of individual subjects to the goggle frames. Diagonal torso straps can be used to prevent movement of the subject's trunk during the experiment if so desired.

The experimental apparatus is mounted in a large rectangular shaped box that is partially open at the subject's end. The other end of the box has doors that can be opened, so that the experimenter can present any desired stimulus combinations and have either the teledeltos surface in position for measurement trials or the lower pins raised for adaptation trials. Indirect lighting is provided by small incandescent bulbs mounted in the top portion of the apparatus.

This apparatus can serve as a tool for investigating sensory-motor adaptation to visual, auditory, or proprioceptive rearrangement or any desired

---

<sup>1</sup> This technique for measuring response position was introduced by Bauer, Woods, and Held (2).

combination thereof. It has already proved valuable in helping to elucidate those sensory factors that must be present before adaptation to visual rearrangement can occur (4).<sup>2</sup>

## REFERENCES

1. BAILY, J. S. Adaptation to prisms: Do proprioceptive changes mediate adapted behaviour with ballistic arm movements? *Quart. J. Exper. Psychol.*, 1972, **24**, 8-20.
2. BAUER, J., WOODS, G., & HELD, R. A device for rapid recording of positioning responses in two dimensions. *Behav. Res. Meth. Instrum.*, 1969, **1**, 157-159.
3. FREEDMAN, S. J., HALL, S. B., & REKOSH, J. H. Effects on hand-eye coordination of two different arm motions during adaptation to displaced vision. *Percept. & Motor Skills*, 1965, **20**, 1054-1056.
4. LACKNER, J. R. Adaptation to displaced vision: Role of proprioception. *Forthcoming*.
5. MIKAEELIAN, H. H. Adaptation to rearranged eye-foot coordination. *Percept. & Psychophys.*, 1970, **8**(4), 222-224.

*Department of Psychology*

*Brandeis University*

*Waltham, Massachusetts 02154*

---

<sup>2</sup> I thank Mr. Arthur Larsen who built the apparatus and Mr. Barkev Bablouzian who developed the recording portion of the apparatus.



## MODELING AND PERSONAL SPACE BEHAVIOR IN CHILDREN\*

*Virginia Commonwealth University*

KENT G. BAILEY, JOHN J. HARTNETT, AND HILDA W. GLOVER<sup>1</sup>

### SUMMARY

Approach *vs.* be approached behavioral measures of personal space were taken on 102 fifth and sixth grade children, each of whom was assigned to one of three groups: (a) Model-Close condition, (b) Model Far condition, and (c) No-Model Control group. A male peer served as model (*M*), and a 41-year-old female served as the object person. Results revealed a strong modeling tendency with both girls and boys tending to stay close or far from the object person as a function of *M* behavior. Boys and girls tended to behave similarly in the Close and Far modeling groups, but girls used more space in the No-Model Control condition. These findings suggest that modeling had an attenuating effect on sex differences in use of space. It was concluded that modeling theory is a viable conceptual tool for use in personal space research.

### A. INTRODUCTION

Following the seminal work of Hall, Little, Sommer, and Keuthe in America, and the detailed analyses of territoriality provided by European ethology, there has developed a rapidly expanding literature in human distance behavior. And, although late in coming, a number of studies have dealt with personal space in children. Personal space development in children has proved to be a favored topic (4, 6, 18, 20, 22), although space dynamics in families (9, 10), mother-child space contingencies (24), social schemata in disturbed adolescents (23), and relationships between acquaintance, liking, and interfigure distances (13) have received attention. All of these studies, however, have dealt exclusively with psychometric, nonbehavioral measures of personal space, including such measures as the Comfortable Interpersonal Distance Scale (4, 5) and a variety of doll,

\* Received in the Editorial Office on July 9, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The third author is now at the Texas Rehabilitation Commission, Houston, Texas. The address of the first two authors is shown at the end of this article.

silhouette, and figure placement tasks. Actual behavioral analyses of personal space are limited to Fry and Willis' (8) study of child-invasion of adult space, and Newman and Pollack's (19) analysis of proxemics in normal *versus* disturbed adolescents. Although Haase and Markey (14) showed that felt board placement and actual behavior were highly correlated, Pedersen (21) found that figure placement and actual behavior were highly correlated for males but were virtually unrelated for females. These issues point to a great need for more actual behavioral studies with both children and adults.

Possibly more serious than the actual-*versus*-simulated behavior problem is the absence of theory in personal space research. It is interesting that the only systematic attempt at theory so far was tailored for use with psychometric as opposed to *in vivo* measures of distance (5). Many writers have lamented the theoretical vacuum that characterizes the area (1, 5, 16), and this lack of theory may explain the methodological and conceptual confusion that now exists. Modeling theory would seem to offer excellent possibilities here, and it is surprising that the ready applicability of modeling principles to proxemic behavior has lain dormant. For example, Flanders (7) reviewed several modeling studies in which dog- or snake-phobic children approached the feared creature more following exposure to a model who experienced no consequences as a result of approach behavior. If children will imitate approach to a feared animal, then it is reasonable to assume that human approach behavior can be manipulated through modeling. Some personal space researchers (1, 3, 15) have used behavioral approach measures with adults that appear to have potential for assessing modeling effects. The appropriate introduction of a model into the approach/be approached technique of Hartnett and Bailey (15) would permit a simple and direct test of sex-of-subject and sex-of-object effects as well as a general modeling effect.

The purpose of the present experiment was to explore the effects of modeling on the personal space behavior of children. On the basis of the general notion of modeling (2, 7), the following specific predictions were made: (a) Ss in the Model Close condition will allow greater proximity to the object person than will Ss in the Model Far condition. (b) Ss in the Model Close condition will allow more proximity with object person than will Ss in a No-Model Control group. (c) Ss in the Model Far condition will stay farther from the object person than will Ss in a No-Model Control group.



## B. METHOD

### 1. Subjects

Subjects were fifth and sixth graders from a single school in the public school system of a moderately large southern city (population 486,000). Except for sex, Ss were randomly assigned to either the Model Far condition (18 males, 18 females) or the Model Close condition (18 males, 18 females). Also selected from the same pool were 15 males and 15 females who were placed in a No-Model group for determining baseline expectations for use of space. This group also served as a control in certain of the analyses. The mean ages per group ranged from 11.49 to 11.57 years. *M* was a male peer, age 11.9, who was described by two teachers as the "most popular" boy in the sixth grade. Although *M* was known to the Ss, none came from his class. The object person was a 41-year-old psychology major who volunteered her help.

### 2. Personal Space Measures

All measures were taken in an experimental room, 32 by 32.6 feet, which was empty except for *S*, *M*, *E*, and the object person. Two-foot strips of tape were placed 18 inches from and parallel to the opposite walls of the room thereby leaving a constant 29.6 feet between starting points.

Two personal space measures were used which have been described in detail elsewhere (1, 15). In the SMI (subject movement index) *S* approaches the object person, while in the EMI (experimenter movement index) the object person approaches the *S*. Following each movement, *E* placed a small marker at the stopping point and, after *S*'s departure, tabulated the distances in tenths of inches.

### 3. Procedures

Preliminary testing with the No-Model group produced an overall mean EMI-SMI approach distance of 8.732 feet, and 8.5 feet was chosen as the standard for defining the experimental conditions. A distance of 6 feet was arbitrarily chosen, so that modeled distances would be 6 feet greater (14.5 feet = Far condition) or 6 feet less (2.5 feet = Close condition) than the standard of 8.5 feet. In the Far condition, the *M* was instructed to approach or be approached at the predetermined distance of 14.5 feet, while in the Close condition he maintained the constant 2.5 feet between himself and the object person. By a prearranged signal from *E* (a small card with

printed word "CLOSE" or "FAR"), *M* was informed as to what condition was appropriate for each new *S*.

When *S* and *M* were escorted into the experimental room, they were given the following general instructions: "We want to find out something about what happens when boys and girls your age meet a stranger. To do this, I'll ask each of you to perform a couple of simple tasks for me. Listen carefully while I tell you what I want you to do. First, will you both come over here and place your toes on this line." Following these instructions, *M* was instructed to walk as close to the object person as he "wanted" to, and then stop. Or, depending on the movement condition, *M* was to allow the object person to approach him until he "wanted" to tell her to stop. *S* followed *M* after each movement and received basically the same instructions.

#### 4. Data Analysis

The main analysis was a mixed  $2 \times 2 \times 2$  factorial design with two between-*Ss* variables (sex and Close-Far modeling condition) and one within-*Ss* variable (type of movement). Since the No-Model baseline group was appropriate to use as a control, two additional analyses (Close condition *vs.* Control and Far condition *vs.* Control) were performed. The  $2 \times 2 \times 2$  designs in the latter two analyses were identical to that used in the main analysis, except that the computer programs were modified to allow for unequal but proportional *ns*.

### C. RESULTS

The mean personal space scores upon which the following analyses are based can be seen in Table 1.

#### 1. Close *vs.* Far Modeling Conditions

Very strong support for hypothesis (a) is provided by the  $2 \times 2 \times 2$  analysis of variance dealing with the Close *versus* Far modeling conditions. The only significant result was on the Close-Far between-*Ss* main effect,

TABLE 1  
MEAN PERSONAL SPACE SCORES (IN FEET) FOR THREE CONDITIONS

Subjects	Model Close			Model Far			Control		
	SMI	EMI	<i>d</i>	SMI	EMI	<i>d</i>	SMI	EMI	<i>d</i>
Boys	3.19	2.84	— .35	11.59	13.02	1.43	6.85	7.81	.96
Girls	4.54	5.10	.56	11.45	13.30	1.85	8.91	11.46	2.55
Total	7.73	7.94	.21	23.04	26.32	3.28	15.76	19.27	3.51

Note: SMI = subject movement index; EMI = experimenter movement index.

but the finding was a very strong one ( $F = 200.37, p < .001$ ). Inspection of the table of means shows the groups differed greatly as a function of  $M$ 's behavior, the Close group using much less space than the Far group. This differential use of space was present regardless of whether the  $S$ 's matching behavior was verbal (saying "Stop" in the EMI condition) or physical (stopping actual approach toward object person in SMI condition).

There was a slight tendency for females to stay farther from the object person in the Close condition than did males while, by contrast, the sexes achieved virtually identical scores in the Far condition. Little can be made of these results, however, in view of the fact that the sex main effect ( $F = 2.48, p > .05$ ) and the sex  $\times$  Close-Far interaction ( $F = 2.12, p > .05$ ) failed to reach significance. The EMI-SMI within- $S$ s approach variable also approached but failed to reach significance ( $F = 2.78, p > .05$ ), and as can be seen, only the Far condition showed a perceptible difference on this variable. In the Far condition both males and females tended to approach the object person to a slightly greater degree than they let the object person approach them.

### 2. Close Condition vs. Control Group

The  $2 \times 2 \times 2$  analysis on the Control *versus* Close groups produced significant between-group main effects on the sex of subject ( $F = 8.01, p < .01$ ) and modeling variable ( $F = 34.62, p < .01$ ), and a significant within-group main effect on the type of movement condition ( $F = 4.40, p < .05$ ). Also the modeling  $\times$  type of movement interaction was significant ( $F = 3.99, p < .05$ ) and proved to be the only interaction to approach significance in the study. The table of means shows that the sex of subject effect is clearly attributable to girls keeping greater distance than boys, and the interaction seems mainly to reflect the much larger EMI-SMI difference in the Control group. Also, as was expected, a strong modeling effect was evident in the Close condition (e.g., the total distance of the combined Control means is more than twice that of the combined Close means). This latter finding provides substantial support for hypothesis (b).

### 3. Far Condition vs. Control Group

The  $2 \times 2 \times 2$  Control *versus* Far analysis of variance yielded a significant modeling main effect ( $F = 17.01, p < .01$ ), thereby supporting hypothesis (c) and a significant type of movement effect ( $F = 7.30, p < .01$ ). The nonsignificant sex variable ( $F = 2.44, p > .05$ ) suggests that the Far modeling condition was less potent in producing sex differences than

was the Close modeling condition. There was a tendency for sex effects to be greatly reduced in the Far condition and were, in fact, virtually nonexistent. Sex differences were, indeed, more obvious in the Control group than in either the Far or Close groups, suggesting that modeling may have served to attenuate rather than augment sex-role characteristics.

#### D. DISCUSSION

The results of the present study provide strong support for the notion that personal space behavior can be manipulated via the principles of modeling. In confirmation of the first three hypotheses, greater proximity was seen in the Model Close *versus* Model Far condition, greater proximity in the Model Close *versus* the No-Model Control condition, and greater distance in the Model Far *versus* the No-Model Control condition. Along with the strong modeling effect, there were some interesting, though not very substantial, sex differences in use of space. In the Close condition, boys did come nearer to *M*'s 2.5 feet than did girls, but this could have been due to the greater social reinforcement value of the female object person for males (11, 12) or, more likely, the greater threat response of females in a personal space situation (1, 4, 17, 18). That threat was a significant variable is suggested by the fact that sex differences virtually disappeared in the presumably less threatening Far condition. However, in view of the finding that the sex differential was greater in the Control than in the presumably more threatening Close condition, modeling may have served to attenuate sex differences up close. Further research with a full matrix of sex variables (male and female models, male and female object persons, etc.) will be needed to clarify these important issues.

One of the most interesting findings from a developmental standpoint was that the approach/be approached variable was the opposite of that normally found in adult college students. Bailey *et al.* (1) and Hartnett *et al.* (15) both reported that subjects allow the object person to come closer than they will approach the object person. However, except for the EMI cell of the male Close group, subjects in the present study tended to approach the object person closer than *vice versa*. This difference is not to be explained by modeling effects, for the largest EMI-SMI difference was found in the female Control group, while the single reversed difference occurred in the cell where males most closely matched *M*'s behavior. Perhaps the adult-child reversal is due to a differential response to the threat of approach, but an artifactual possibility must also be considered: In the present study, the starting distance between subject and object was 29.6 feet,



more than twice that of the prior studies with adults. Actual developmental comparisons, based on identical starting distances, are called for to assess the validity of the apparent reversal obtained in the present study.

There may be some question as to whether the present study addressed itself to modeling or some other social psychological phenomenon, such as "set," "expectation," "conformity," or "social influence." Some of these distinctions are difficult to make, and often the issue is decided by author preference. Bandura (2) allows the term "modeling" to cover a wide range of antecedent, mediating, and behavioral variables heretofore subsumed under a variety of terms. He argues that the same learning process is involved regardless of the generality of what is learned, the model's characteristics, and the stimulus conditions under which matching is performed. Therefore, if *S* significantly modifies his behavior as a function of exposure to *M*'s behavior—regardless of the prevailing conditions—then *S* has "modeled" after *M*. In the present study, *M* performed a clearly discriminable response (saying "Stop" in EMI and stopping at a certain point in SMI) which provided opportunity for modeling to occur. Whether reinforcement was involved or not is unclear, but it may be that *S* attempted to "please" *M*, the experimenter, or both. Again the study provides many more questions than answers, but it seems fair to conclude that modeling theory is a viable conceptual and heuristic device that has broad applicability in personal space research.

#### REFERENCES

1. BAILEY, K. G., HARTNETT, J. J., & GIBSON, F. W., JR. Implied threat and the territorial factor in personal space. *Psychol. Rep.*, 1972, **30**, 263-270.
2. BANDURA, A. *Principles of Behavior Modification*. New York: Holt, Rinehart & Winston, 1969.
3. DOSEY, M., & MEISELS, M. Personal space and self-protection. *J. Personal. & Soc. Psychol.*, 1969, **11**, 93-97.
4. DUKE, M. P., & NOWICKI, S., JR. Perceived interpersonal distance as a function of the subject's locus of control and the race and sex of stimuli in elementary and high school children. Paper presented at the meeting of the Southeastern Psychological Association, Miami, Florida, 1971.
5. DUKE, M., & NOWICKI, S. A new measure and social-learning model for interpersonal distance. *J. Exper. Res. Personal.*, 1972, **6**, 119-132.
6. ESTES, B. W., & RUSH, D. Social schemas: A developmental study. *J. of Psychol.*, 1971, **78**, 119-123.
7. FLANDERS, J. P. A review of research on imitative behavior. *Psychol. Bull.*, 1968, **64**(5), 316-337.
8. FRY, A. M., & WILLIS, F. H. Invasion of personal space as a function of the age of the invader. *Psychol. Rec.*, 1971, **21**, 385-389.
9. GERBER, G. L. Psychological distance in the families of normal, disturbed, and learning-problem children. *J. Consult. & Clin. Psychol.*, 1973, **40**, 139-147.



10. GERBER, G. L., & KASWAN, J. Expression of emotion through family grouping schemata, distance, and interpersonal focus. *J. Consult. & Clin. Psychol.*, 1971, **36**, 370-377.
11. GEWIRTZ, J. L. Three determinants of attention-seeking in young children. *Monog. Soc. Res. Child Devel.*, 1954, **19**(2), No. 59.
12. GEWIRTZ, J. L., & BEAR, D. M. The effect of brief social deprivation on behaviors for a social reinforcer. *J. Abn. & Soc. Psychol.*, 1958, **56**, 49-56.
13. GUARDO, C. Personal space in children. *Child Devel.*, 1969, **40**, 143-151.
14. HAASE, R. F., & MARKEY, M. J. A methodological note on the study of personal space. *J. Consult. & Clin. Psychol.*, 1973, **40**, 122-125.
15. HARTNETT, J. J., BAILEY, K. G., & GIBSON, F. W., JR. Personal space as influenced by sex and type of movement. *J. of Psychol.*, 1970, **76**, 139-144.
16. LINDER, D. E. Too close for comfort. *Contemp. Psychol.*, 1970, **15**, 40-41.
17. LITTLE, K. B. Personal space. *J. Exper Soc. Psychol.*, 1965, **1**, 237-247.
18. MEISELS, M., & GUARDO, C. J. Development of personal space schemata. *Child Devel.*, 1969, **40**, 1166-1177.
19. NEWMAN, R. C., & POLLACK, D. Proxemics in deviant adolescents. *J. Consult & Clin. Psychol.*, 1973, **40**, 6-8.
20. PEDERSEN, D. M. Developmental trends in personal space. *J. of Psychol.*, 1973, **83**, 3-9.
21. ———. Relations among sensation seeking and simulated and behavioral personal space. *J. of Psychol.*, 1973, **83**, 79-88.
22. TOLOR, A. Psychological distance in disturbed and normal children. *Psychol. Rep.*, 1968, **23**, 695-701.
23. WEINSTEIN, L. Social schemata of emotionally disturbed boys. *J. Abn. & Soc. Psychol.*, 1965, **70**, 457-461.
24. ———. The mother-child schema, anxiety, and academic achievement in elementary school boys. *Child Devel.*, 1968, **39**, 257-264.

*Department of Psychology*  
*Virginia Commonwealth University*  
*800 West Franklin Street*  
*Richmond, Virginia 23220*

## PERSONALITY CORRELATES OF PROXIMITY PREFERENCES\*

*Departments of Psychology, DePaul University and Illinois Institute of Technology*

---

ALAN F. SEWELL<sup>1</sup> AND JAMES T. HEISLER

---

### SUMMARY

Each of 35 male Ss entered a room and was greeted by a male interviewer seated at the opposite end of the room, who asked him to pull up a chair. A measurement was taken of the distance at which S positioned his chair from E's (proximity measurement). Following a brief orientation on the alleged purpose of the experiment, S was given the Personality Research Form (PRF) and asked to complete it at home.

Correlations were computed between scores on the 22 scales of the PRF and proximity measurements. Significant negative correlations were obtained between proximity measures and scores on the "Exhibition" scale ( $r = -.43$ ,  $p < .01$ ) and the "Impulsivity" scale ( $r = -.36$ ,  $p < .05$ ). The reliability of such proximity measures is reported from a previous study. The results are discussed in terms of Hall's classification system of interpersonal space.

---

### A. INTRODUCTION

The relatively recent consideration of physical space as an important variable in human interaction has generated numerous scattered studies of spatial factors in social interactions. Two excellent reviews, one by Sommer (8) and the other by Patterson (5), illustrate the diversity of the work done in this area. Hall (1) has contributed theoretical notions to this area of psychological investigation in his discussions of the communicative function of varying interpersonal distances. Spatial arrangements between parties are believed to be highly structured by cultural norms. But even within cultures, there is often a great deal of variability in the spatial behavior of interacting individuals. Perhaps the interactants themselves account for this intracultural variability. More precisely, perhaps personality characteristics account for at least some of this dissimilarity in spatial behavior.

---

\* Received in the Editorial Office on July 9, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> Reprints are available from the first author at the address shown at the end of this article.

Relatively few studies have attempted to isolate or identify personality factors that may relate to physical proximity in interpersonal relations. The work that has been done in relating personality variables to spatial interaction typically has been inconclusive. Indicative of these inconclusive findings has been the work with the personality characteristics of extroversion and introversion. Leipold (4) and Patterson and Holmes (6) reported that an extrovert would approach an interviewer more closely and talk longer in response to questions than would an introvert. However, Williams (9) found that there was no difference in the approach tendencies of extroverts and introverts. In the same study, Williams did, however, find that extroverts would allow others to approach them more closely than would introverts.

Thus the relatively few investigations into the relationship between personality characteristics and spatial behavior have failed to yield definitive data. Moreover, there appears to have been no specific documentation of the reliability of proximity as a personal variable. The purpose of the present study was to explore interrelations between personality characteristics and proximity preferences. This article also reports the findings of an earlier investigation by Sewell and Heisler (7) which investigated the reliability of proximity as a personal variable.

## B. METHOD

### 1. Subjects

The Ss were 35 male undergraduates at DePaul University, whose ages ranged from 18 to 22 years and who were nonvolunteers from Introductory Psychology classes.

### 2. Apparatus

The experiment was conducted in a rectangular interview room (22 feet 4 inches  $\times$  17 feet  $\times$  12 feet 6 inches). A male interviewer (*E*), who was the same person for all subjects, was seated in front of a desk at the far end of the room. Two chairs were positioned to the left and right of the doorway which was at the opposite end of the room from where *E* was seated.

Each *S* came to the room, ostensibly for orientation to a research project dealing with personality characteristics of today's undergraduates as compared to those of 20 years ago. Upon entering the room, Ss were greeted by the seated *E* who said, "Hello, please pull up a chair." The distance from *E* at which each *S* positioned his chair was the proximity measure (the distance between the centers of *S*'s chair and *E*'s chair). *E* then instructed each *S* as follows:

I am doing a study in which I am comparing certain personality characteristics of today's college undergraduates with undergraduates of 20 years ago. I would like you to take this test home and complete it in your spare time and return it to the Psychology Office as soon as possible. Do you have any questions?

*E* handed each *S* a copy of Jackson's (2) Personality Research Form (Form A) and explained how to mark the answer sheet. *E* then thanked *S* for his participation. *S* then was dismissed. The true purpose of the experiment was revealed to *S* at the time he returned his completed Personality Research Form.

In an earlier and similar study (7) two proximity measures were obtained. Upon entering the room, *S* was told to pull up a chair. The distance from *E* at which *S* positioned his chair was Proximity Measure #1. *S* was then sent from the room on an errand related to the experiment, and *E* also left the room. During their absence, *S*'s chair was returned to its original position. Upon returning, *S* was again asked to seat himself, thereby providing Proximity Measure #2.

### C. RESULTS

In the original study (7) the reliability of proximity as a personal variable was demonstrated ( $r = .96$ ,  $p < .001$ ). The present study assumed this reliability; thus the reason for only one proximity measure. In this study, proximity measures ranged from 33.5 inches to 206 inches. The median distance was 151.5 inches, and the mean distance was 135.96 inches. Pearson  $r$ s were computed between proximity measures and scores for each of the 22 scales of the Personality Research Form.

The results are as follows: "Abasement,"  $-.05$ ; "Achievement,"  $-.04$ ; "Affiliation,"  $-.16$ ; "Aggression,"  $-.28$ ; "Autonomy,"  $-.13$ ; "Change,"  $-.28$ ; "Cognitive Structure,"  $-.31$ ; "Defendence,"  $.06$ ; "Dominance,"  $-.20$ ; "Endurance,"  $.11$ ; "Exhibition,"  $-.43$ ; "Harmavoidance,"  $.05$ ; "Impulsivity,"  $-.36$ ; "Nurturance,"  $-.17$ ; "Order,"  $.13$ ; "Play,"  $-.07$ ; "Sentience,"  $-.05$ ; "Social Recognition,"  $-.21$ ; "Succorance,"  $-.05$ ; "Understanding,"  $-.16$ ; "Desirability,"  $-.04$ ; "Infrequency,"  $-.14$ . Two of these  $r$ s are significant: "Exhibition" ( $r = -.43$ ,  $p < .01$ ) and "Impulsivity" ( $r = -.36$ ,  $p < .05$ ).

### D. DISCUSSION

The major significant finding in this study was that people who scored high on the "Exhibition" scale tended to position themselves relatively close

to *E*. This finding seems reasonable in view of Jackson's (3, p. 6) description of high scorers: "They want to be the center of attention; they enjoy having an audience; they engage in behavior which wins the notice of others." Jackson offers the following trait-defining adjectives: "conspicuous, expressive, demonstrative, and pretentious." Subjects who scored high on the "Impulsivity" scale also tended to position themselves relatively close to *E*. This, too, seems reasonable in that Jackson defines high scorers on this scale as people who "tend to act on the 'spur of the moment' and without deliberation; who give vent readily to feelings and wishes; who speak freely; and who may be volatile in emotional expression" (p. 7). Not statistically significant but of conceptual importance was the finding that high scorers on the "Cognitive Structure" scale tended to sit relatively far from *E*. According to Jackson, high scorers on this scale are people who "do not like ambiguity or uncertainty in information" (p. 6). Jackson describes high scorers on this scale with such terms as "rigid," "perfectionistic," and "need structure." This trait appears to be somewhat juxtaposed to what Jackson calls "Impulsivity." In fact, normative data collected by Jackson show that all three scales—"Exhibition," "Impulsivity," and "Cognitive Structure"—are intercorrelated. "Cognitive Structure" is negatively correlated with the other two scales, which are positively correlated with each other.

From a theoretical standpoint, these findings are in line with Hall's (1) notion of the communicative function of interpersonal space. Specifically, *Ss* who scored high on "Exhibition" and "Impulsivity"—*Ss* who might be described as outward going or extraverted—tended to position themselves anywhere from three to five feet from *E*, in contrast to other *Ss* (e.g., high scorers on "Cognitive Structure" who, having been low scorers on "Exhibition" and "Impulsivity," might be described as socially cautious) who sat at greater distances. Interpersonal distances of three to five feet bridge two distances in Hall's classification system of interpersonal space—Personal Distance (Far Phase) and Social Distance (Close Phase). According to Hall, there is a sense of physical and psychological involvement at these distances (three to five feet). Moreover, approaching a person at these relatively close distances can communicate to that person a sense of dominance on the part of the approaching individual or, at the very least, a feeling that the two interactants are on equal footing. On the other hand, at greater distances, there is decreasing physical and psychological involvement, making such distances "psychologically safe."

In terms of earlier research in this general area, the findings of this study



tend to support the findings of both Leipold (4) and Patterson and Holmes (6) who reported that extroverts would approach an interviewer more closely.

## REFERENCES

1. HALL, E. T. *The Hidden Dimension*. New York: Doubleday, 1966.
2. JACKSON, D. N. *Personality Research Form: Form A*. University of Western Ontario, London, Canada, 1964.
3. ———. *Personality Research Form Manual*. Goshen, N. Y.: Res. Psychol., 1967.
4. LEIPOLD, W. D. Psychological distance in a dyadic interview as a function of introversion-extroversion, anxiety, social desirability, and stress. Unpublished Doctoral dissertation, University of North Dakota, Grand Forks, 1963.
5. PATTERSON, M. L. Spatial factors in social interactions. *Hum. Relat.*, 1968, **21**, 351-361.
6. PATTERSON, M. L., & HOLMES, O. S. Social interaction correlates of the MMPI extroversion-introversion scale. *Amer. Psychol.*, 1966, **21**, 724-725 (abst.).
7. SEWELL, A. F., & HEISLER, J. T. Personality and proximity. Unpublished study, DePaul University, Chicago, Illinois, 1969.
8. SOMMER, R. Small group ecology. *Psychol. Bull.*, 1967, **67**, 145-152.
9. WILLIAMS, J. L. Personal space and its relation to extroversion-introversion. Unpublished Master's thesis, University of Alberta, Edmonton, Canada, 1963. Cited by R. Sommer (8).

*Department of Psychology*  
*DePaul University*  
*2219 North Kenmore Avenue*  
*Chicago, Illinois 60614*



## PERSON PERCEPTION AS A FUNCTION OF THE PERSONAL CONSEQUENCES AND IMMEDIACY OF A DECISION\*

*Department of Psychology, DePaul University*

---

ALAN F. SEWELL<sup>1</sup>

---

### SUMMARY

The hypothesis that situational variables differentially influence the direction and magnitude of person perception was tested in this study. Ss read identical biographical descriptions of two fictitious graduate assistants under one of three conditions: (a) instructions established no personal consequences, (b) instructions required S to choose the assistant who would write and grade an examination, (c) instructions required S to indicate a preference for one of the assistants although an assignment had already been made. S then rated each of the two assistants on Evaluation and Potency semantic differential scales. Sharply differentiated ratings occurred under the first condition and, to a lesser degree, under the third. Under the second condition ratings of the assistants were not significantly different, suggesting an "error cost" conservatism in differential perception when a decision has personal consequences. It was concluded that situational demands may be more potent determinants of person perception than either the characteristics of the stimulus person or the personality characteristics of the perceiver.

---

### A. INTRODUCTION

Person perception is one of the more active areas of research in social psychology. Most of the generated research, however, has been concerned with the manner in which an impression of a person (typically either fictitious or unknown to the subject) is formed. Hastorf *et al.* (5, pp. 35-60) have provided an excellent review of the literature in this area of investigation; as these authors quite correctly observe "... most recent research in person perception has focused on how information about another is processed" (p. 35).

The "how" of such studies is, of course, quite closely related to the "what"

---

\* Received in the Editorial Office on July 9, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup> The author is grateful for the assistance of Dr. R. E. Brewer in some aspects of the design of this research and some preliminary studies.

of person perception; that is, the manner in which information about another is processed quite probably is dependent upon the kinds of information available. In a similar way two other aspects of person perception are also quite probably interrelated: the "why" and "when" of person perception. These aspects of perception are relatively indifferent to modes and models of information processing; they are, instead, concerned with the circumstances that lead to specific and reliable perceptions of others.

It is possible to conceive of a need to evaluate others as stemming from some underlying personality characteristic; the paranoid, for example, in constant expectation of attack, may need to evaluate the risk potential of strangers. It seems more reasonable, however, to hypothesize that most people form impressions of others only as some circumstantial need requires. Hence, situational determinants of person perception would be more powerful than personal determinants, and situational demands would occasion differential perceptions of persons.

Research in the area of situational demands for differential perceptions is scanty. Darley and Berscheid (2) found that subjects who anticipated interaction with a "partner" liked that person better than a "non-partner," but they generally concluded that differential evaluation, in the form of over-evaluation of the "partner," was not a significant means of rationalizing the greater liking. Berscheid *et al.* (1), however, did find overevaluation of significant characteristics of the anticipated "partner." These studies, however, appear to have employed a relatively powerful impetus to differential perceptions: subjects expected to engage in extensive and detailed discussions of highly personal matters. The clearest conclusion to be derived from these studies is that the anticipation of interaction is a powerful determinant of differential person perceptions, a conclusion widely reported in textbooks of social psychology.

The present study postulates the existence of a continuum of situational demands which will differentially influence person perceptions. In a preliminary study, subjects were found to generate sharply differential evaluations of opposing participants in a fictitious lawsuit, significantly lower evaluations being accorded to the participant judged "guilty," despite the fact that no evidence was adduced for or against either participant (9). To test the hypothesis that the subject's commitment (in the form of the "verdict") was a critical factor, a second study required subjects to evaluate the instructor of a course in which they were registered and another instructor who simply made an announcement on the first day of class: surprisingly, the

ratings of these two persons did not differ significantly despite the fact that course registration would appear to constitute a commitment.

The present study was designed specifically to test a hypothesis generated to account for the results of the preceding study: that while the subject's commitment was valid, it had not been activated by specific situational demands. That is, the subject had not been required to recommit himself by making a decision or by indicating a personal preference. In the terminology devised by Heider (6), the preceding study had not enforced a "unit relationship" for the subject; in the conceptual framework proposed by Jones and Davis (7), the rating task was deficient in "hedonic relevance" (pp. 237-240). The present study attempted to vary the relevance and immediacy of the imposed situational demands.

## B. METHOD

### 1. *Materials*

Three types of booklets were prepared, each intended to represent a different position along a continuum of situational demands. Each booklet consisted of seven pages, the last six of which were nearly identical in all booklet types; the chief difference between the booklet types was the nature of the instructions contained on the first page.

Instructions for completion of Type 1 booklets were intended to establish *minimal* situational demands. Subjects were merely asked to evaluate two persons as part of a study of impression formation based upon limited biographical information.

Instructions for completing Type 2 booklets were intended to establish *maximal* situational demands; subjects were informed they were participating in an educational evaluation project which required each student to take a midterm examination prepared by one of two graduate assistants. Each student was to choose the assistant whose examination he preferred and then to evaluate both assistants.

Instructions for Type 3 booklets attempted to establish an intermediate level of situational demands. Subjects were again informed they would receive midterm examinations prepared by two graduate assistants. In this case, however, students were not allowed to select the assistant: each student had been randomly assigned to one of the two assistants. Each student was informed to which assistant he had been assigned, but was asked to indicate a preference nonetheless and then to rate the two assistants.



Following the instructions page in all booklet types were the two pages containing descriptions of the fictitious graduate students, as follows. The order of presentation of these descriptions was counterbalanced within each booklet type.

David T. is a first-year graduate student at DePaul. He is 23 years old and unmarried. He is 6 feet, 1 inch tall; he has light brown hair and brown eyes; and weighs 170 pounds. David received a Bachelor's degree in psychology from William Penn College in Philadelphia. He hopes to earn a doctorate in clinical psychology. He says he especially enjoys working with children and young adults because he is more familiar with their problems.

Robert D. is a 24-year-old graduate student at DePaul. He is single; 6 feet, 2 inches tall; weighs 185 pounds; and has blond hair and blue eyes. He received his undergraduate degree from Southern Minnesota State University, majoring in psychology. Now a first-year graduate student of clinical psychology, Robert is more interested in research than in clinical practice. He believes that careful and intensive research is critically important to progress in psychology.

The page following the descriptions required each *S* to indicate a preference for one of the two graduate assistants as preparer of his midterm examination and then to indicate the confidence of this decision along a four-point scale. This page appeared in all booklet types.

The next page contained instructions for completing semantic differential scales; the instructions were derived from those suggested by Osgood *et al.* (8). The remaining two pages of all booklets contained a series of six semantic differential scales to be used in evaluating the two graduate assistants. Each page was headed with the name of one of the assistants. The six scales were employed in the following order: *good-bad*, *foolish-wise*, *severe-lenient*, *hard-soft*, *kind-cruel*, *strong-weak*. This order had been randomly chosen, but was identical for all booklets. These scales were chosen to secure ratings on two different and poorly correlated factors identified by Osgood *et al.*: an Evaluation factor (*good-bad*, *foolish-wise*, *kind-cruel*) and a Potency factor (*severe-lenient*, *hard-soft*, *strong-weak*). The three selected scales were those that had been shown to load most heavily on the factor.

## 2. Subjects and Procedure

The booklets were randomly distributed in a class in Introductory Psychology. Nineteen *Ss* completed Type 1 booklets, 12 completed Type 2 booklets, and 16 completed Type 3 booklets. Male and female *Ss* were approximately equally represented in all groups. Data were collected during the class period immediately preceding the scheduled date of the midterm

examination. The purpose of the study was explained after data had been collected.

### C. RESULTS

A preliminary study (9) had shown no differences in results regardless of the order of decision and rating; hence no analysis of order effects was undertaken here. Separate analyses of Evaluation and Potency ratings were performed on the basis of the low correlation of these factors (8).

Each scale was scored from 1 to 5, the higher number indicating the more positive pole ("good" or "strong"). For each *S* a mean factor score was calculated for each rated person; hence each *S* provided four scores, two of which applied to the "Preferred" (P) person, and two of which applied to the "Nonpreferred" (N-P) person.

Mean Evaluation ratings were as follows: in the Type 1 situation,  $P = 3.98$ ,  $N-P = 3.40$ ; Type 2,  $P = 3.56$ ,  $N-P = 3.44$ ; Type 3,  $P = 3.69$ ,  $N-P = 3.46$ . Mean Potency ratings were as follows: Type 1,  $P = 2.79$ ,  $N-P = 3.72$ ; Type 2,  $P = 3.03$ ,  $N-P = 3.18$ ; Type 3,  $P = 3.10$ ,  $N-P = 3.46$ .

Since the primary focus of this study was upon differential perceptions under the three situational demand conditions, analyses of variance were performed upon differences in ratings of the P and N-P persons. Two separate one-way analyses were performed.

For the Evaluation analysis, the obtained  $F = 7.150$  ( $df = 2/45$ ,  $p < .01$ ). For the Potency analysis, the obtained  $F = 20.568$  ( $df = 2/45$ ,  $p < .001$ ). *Post-hoc* comparisons showed significant differences between Evaluation mean differences for Type 1 and Type 2 ( $t = 3.822$ ,  $df = 30$ ,  $p < .001$ ) and between Type 1 and Type 3 ( $t = 2.692$ ,  $df = 33$ ,  $p < .05$ ); Type 2 and Type 3 did not differ significantly ( $t = .731$ ,  $df = 27$ ,  $p > .05$ ). For the Potency mean differences, *post-hoc* comparisons were all statistically significant: Type 1-Type 2 ( $t = 4.060$ ,  $df = 30$ ,  $p < .001$ ); Type 2-Type 3 ( $t = 2.642$ ,  $df = 27$ ,  $p < .05$ ); and Type 1-Type 3 ( $t = 6.751$ ,  $df = 33$ ,  $p < .001$ ). Thus the instructions that established situational demand characteristics were effective in producing different patterns of person perception and evaluations.

A subsequent analysis was concerned with differences in ratings of P and N-P within each group. For the Evaluation factor, these differences were significant in the Type 1 and Type 3 situations: Type 1,  $t = 19.300$ ,  $df = 18$ ,  $p < .001$ ; Type 2,  $t = 1.641$ ,  $df = 11$ ,  $p > .05$ ; Type 3,  $t = 2.181$ ,  $df = 15$ ,  $p < .05$ . The same pattern was obtained with respect to the Potency

factor: Type 1,  $t = 5.882$ ,  $p < .001$ ; Type 2,  $t = 1.230$ ,  $p > .05$ ; Type 3,  $t = 2.572$ ,  $p < .05$ .

#### D. DISCUSSION

The results of this study have been graphed in Figure 1. The manner in which this graph has been constructed implies, through interconnection of the groups, that a process continuum underlies differential person perceptions: the process is determined by situational demands that can be characterized by the relevance of a decision and by the immediacy of the consequences of the decision.

The Type 1 situation actually required no decision—simply a statement of preference. The distinction between a “decision” and a “statement of preference” has been emphasized by Festinger (3, p. 156): a decision has behavioral consequences, while a preference typically has none. Ratings are essentially statements of preferences, and they remain simply that unless

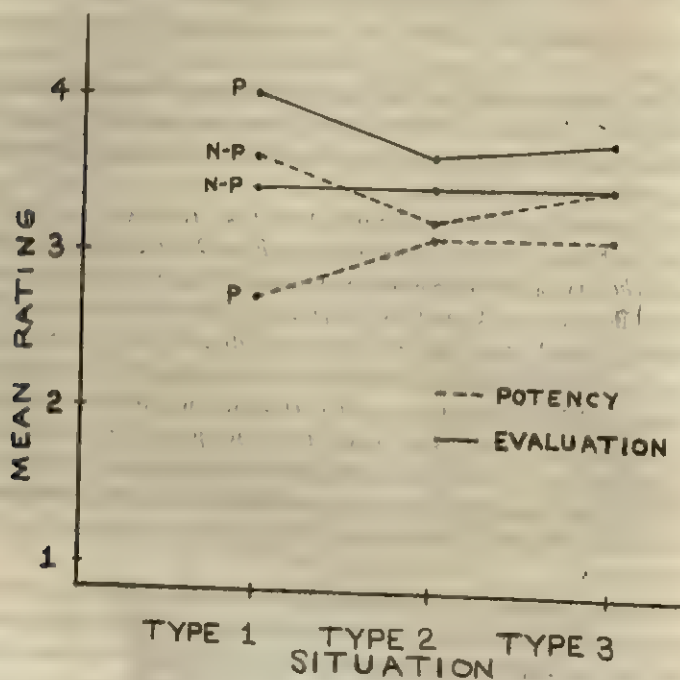


FIGURE 1

HYPOTHESIZED PROCESS CONTINUUM OF DIFFERENTIAL PERSON PERCEPTION AS  
A FUNCTION OF SITUATIONAL DEMANDS

Data points plotted represent results obtained in the present study.

they are associated with a true decision. The Type 2 situation employed in this study does require a genuine decision, the consequences of which are both personal and immediate. The Type 3 situation employed here evidently is intermediate: the subject's choice is something more than a simple preference, but something less than an effective decision.

As Figure 1 shows, a surprising effect was encountered in this study. Sharply differentiated perceptions appear easy under Type 1 and Type 3 situational demands, but very difficult under the demands of Type 2 situations. A possible explanation of this phenomenon can be provided by a concept of "error cost": misperceptions in Type 1 situations are inexpensive in that they lack consequences. Similarly, misperceptions in Type 3 situations are relatively inexpensive, since the subject's decision-preference is unlikely to affect a predetermined course of events. Misperceptions in Type 2 situations, however, can be very costly in terms of behavioral consequences; hence the subject responds conservatively.

Another phenomenon evident in Figure 1 has been consistently recorded in preliminary studies: the "liked" person is not the "powerful" person; that is, while P receives high Evaluation factor ratings, he receives low Potency ratings, and the reverse is true of N-P. This phenomenon appears regardless of the situation type. Hence it would seem to be more closely related to simple preferences than to decisional preferences.

The biographical descriptions of "David" and "Robert" were deliberately prepared to permit physical differentiation of the persons on the basis of hair and eye color, height, and weight, but without any expectation that these characteristics would influence preferences in any systematic way. The descriptions of career plans, however, were expected to provide a basis for systematic preferences; "David" would appear to be more sympathetic to students' needs and problems. When preferences in Type 1 and Type 2 situations are counted, this expectation is confirmed: "David" was the unanimous choice in the Type 1 situation (no decision) and the choice of 11 of the 13 subjects in the Type 2 situation (requiring an effective decision). In the Type 3 situation (decision without probability of consequences), only 11 of 16 subjects chose "David"; of these 11, nine had been informed that "David" had been assigned to them. More strikingly, "Robert," nonpreferred in Type 1 and Type 2 situations, remained the choice of nearly one-third of the Type 3 subjects; only two of the subjects who had had "Robert" assigned chose "David." These results generally agree with those reported by Berscheid *et al.* (1), except that a substantially greater percentage of "non-switchers" was encountered in the present study. Those investigators found, however, that "non-switchers display more prechoice liking for their partners



than switchers" (p. 18), and this conclusion is clearly not supported by the data of the present study, as indicated by comparative preferences in Type 1, Type 2, and Type 3 situations. Here it is the fact of assignment or freedom of choice that directs the liking.

Three bases for the directionality and strength of differential person perceptions can be hypothesized: the perceived characteristics of the persons to be evaluated, idiosyncratic (personality) characteristics of the perceiver, and the demands of the perceptual situation. Random assignment of subjects to the three situations of this study eliminated the second of these possibilities, while the methodology of the study permits an evaluation of the relative influence of each of the other possibilities. The results of the study provide strong evidence that identical persons can be very differently perceived under different situational demand conditions. It is the situation that directionalizes and determines the strengths of perceptions far more effectively than the characteristics of the stimulus persons.

Further investigation of situational influences upon social psychological phenomena appears clearly to be in order. Indeed, recommendations for such investigations were recently published by Fredericksen (4), and the present study suggests at least one methodological approach to such investigations.

#### REFERENCES

1. BERSCHIED, E., BOYE, D., & DARLEY, J. M. Effect of forced association upon voluntary choice to associate. *J. Personal. & Soc. Psychol.*, 1968, 8, 13-19.
2. DARLEY, J. M., & BERSCHIED, E. Increased liking as a result of the anticipation of personal contact. *Hum. Relat.*, 1967, 20, 29-40.
3. FERTINGER, I. Conflict, Decision, and Dissonance. Stanford, Calif.: Stanford Univ. Press, 1964.
4. FREDERICKSEN, N. Toward a taxonomy of situations. *Amer. Psychol.*, 1972, 27, 114-123.
5. HASTORY, A. H., SCHNEIDER, D. J., & POLEYKA, J. Person Perception. Reading, Mass.: Addison-Wesley, 1970.
6. HEIDER, F. The Psychology of Interpersonal Relations. New York: Wiley, 1958.
7. JONES, E. E., & DAVIS, K. E. From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 2). New York: Academic Press, 1965. Pp. 219-266.
8. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. The Measurement of Meaning. Urbana: Univ. Illinois Press, 1957.
9. SEWELL, A. F. Vocabulary, persuasion, and person perception. Unpublished Doctoral dissertation, DePaul University, Chicago, Illinois, 1971.

Department of Psychology  
DePaul University  
2219 North Kenmore Avenue  
Chicago, Illinois 60614



## THE EPSAT AS A PREDICTOR\*

*Eastern Michigan University*

---

JERRY W. O'DELL

---

### SUMMARY

Two independent studies were done, both indicating that the Engineering and Physical Science Aptitude Test (EPSAT) and the Graduate Record Examination (GRE) offer promise in prediction of performance in graduate statistics courses.

---

### A. INTRODUCTION

In a recent issue of this journal, Kooker (3) reported that he had found little correlation between scores in a graduate statistics course and the Miller Analogies Test. The present writer is familiar with the low correlations that generally obtain between predictors and grades in graduate school: e. g., in Lannholm's (4) survey. However, Kooker's findings were still surprising to the writer, for in the past year he had compared Graduate Record Examination (GRE) scores and scores on the Engineering and Physical Science Aptitude Test (EPSAT) (5) with grades in his graduate course in psychological statistics and had obtained substantial correlations (see the top half of Table 1).

It was possible, however, that the results might be spurious, for the writer had known the EPSAT scores before assigning grades in the course. Therefore, before rushing into print with the findings, it was decided to attempt the study a second year, making sure that no prior knowledge of results was possible. The following is a report of the two studies.

### B. METHOD

Two classes in graduate psychological statistics were used, one in the Fall of 1971, one in the Fall of 1972. Both classes consisted of 31 subjects, virtually all of whom were working toward a master's degree in psychology. The EPSAT mathematical sections (I, II, and IV) were administered in the first class session and scored as indicated in the test manual. Graduate

---

\* Received in the Editorial Office on July 11, 1973, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

TABLE 1  
CORRELATIONS BETWEEN VARIABLES

Variables	EPSAT		GRE-V		GRE-Q		GRE-A		Stat	
	<i>r</i>	<i>N</i>	<i>r</i>	<i>N</i>	<i>r</i>	<i>N</i>	<i>r</i>	<i>N</i>	<i>r</i>	<i>N</i>
EPSAT			.39	23	.79***	23	.18	18	.64***	28
GRE-V	.54**	25			.32	26	.46*	21	.58**	25
GRE-Q	.72***	25	.30	26			.18	21	.46*	25
GRE-A	.63**	19	.84***	19	.49*	19			.27	20
Stat	.43*	30	.21	26	.46*	26	.23	19		

*Note:* Correlations for the 1971 class are in the top half of the table, and those for 1972 are in the bottom. Sample sizes for each correlation are shown in the *N* columns. Definitions of abbreviations are as follows: EPSAT = Engineering and Physical Science Aptitude Test mathematical sections; GRE = Graduate Record Examination (V = Verbal, Q = Quantitative, A = Psychology Aptitude); Stat = statistics class grade.

\* Significant at the .05 level.

\*\* Significant at the .01 level.

\*\*\* Significant at the .001 level.

Record Examination data were obtained from the students' departmental files. The criterion measure was simply the total number of points on examinations taken in the course. The major difference between the two years consisted in the fact that strict procedures were maintained to prevent contamination of assignment of grades by test scores in the second year. Tests were scored by an assistant, who kept them until after the grades were submitted. GRE scores were similarly not looked up until all the grades were in. Simple product-moment correlations were obtained through the use of the Biomedical computer program for missing data correlations, for we did not have complete data on all subjects.

### C. RESULTS AND DISCUSSION

Table 1 contains both years' results. It may be seen that there are strong correlations between the criterion and predictors in both cases, albeit the correlations are rather lower the second year. Fortunately for the writer's peace of mind, however, the difference in correlation between the EPSAT and statistics grades for the two years (.6355 and .4344) was not statistically significant ( $t = 1.03$ ).

Surprisingly, in both studies, the Engineering and Physical Science Aptitude Test, which, if entries in Buros (1) are any indication, is not a highly regarded or widely used test, worked out as a rather effective predictor of success in this graduate course. Further, the GRE scores were also good predictors, especially the quantitative scores.

It is the writer's feeling that the relative success of the EPSAT in predicting

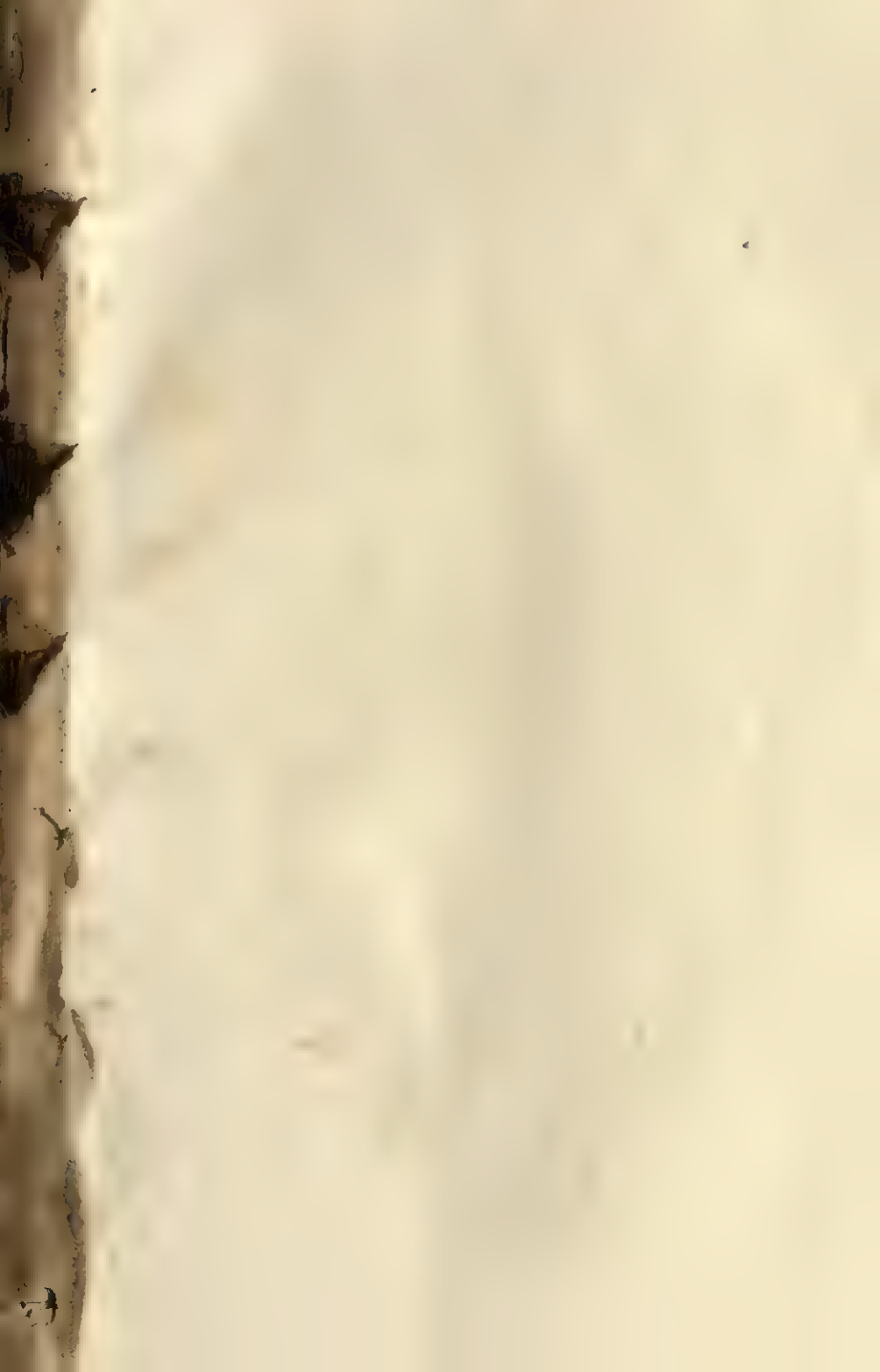
course grades lay in the rationale behind its selection for this study. It was felt that the statistics course is more a course in mathematics than in psychology *per se* and hence could best be predicted by a mathematics proficiency test. The EPSAT simply happened to be the most readily available commercial instrument with a rather difficult mathematics section. In short, it was felt that if the test were tailored to the course and its students, more directly sampling the subject matter than is usually the case, prediction would be optimized. This may also be seen in the fact that the GRE Quantitative section also predicted well, while the Verbal section was less consistent. The results may also say something about the field of psychology in that so little relationship was found between the GRE Psychology scores and course grades! Further, it is unlikely that we are dealing with chance findings here, since the same results were obtained twice. It is also interesting that our correlations are similar to those obtained in early work with the EPSAT. For example, Gregg (2) reports a correlation of  $+.63$  between the test and course grades in engineering.

## REFERENCES

1. BUROS, O. K., Ed. The Seventh Mental Measurements Yearbook. Highland Park, N. J.: Gryphon Press, 1972.
2. GREGG, G. W. An investigation of the reliability and validity of the Engineering and Physical Science Aptitude Test. *J. Educ. Res.*, 1951, **45**, 299-305.
3. KOOKER, E. W. The relationship between performance in a graduate course in statistics and the Miller Analogies Test and the Watson-Glaser Critical Thinking Appraisal. *J. of Psychol.*, 1971, **77**, 165-169.
4. LANNHOLM, G. V. Review of studies employing GRE scores in predicting success in graduate study, 1952-1967. Special Report No. 67-4, Educational Testing Service, Princeton, New Jersey, 1967.
5. MOORE, B. V., LAPP, C. J., & GRIFFIN, C. H. The Engineering and Physical Science Aptitude Test. New York: Psychol. Corp., 1943.

*Psychology Department*  
*Eastern Michigan University*  
*Ypsilanti, Michigan 48197*









## ABBREVIATIONS OF WORDS USED IN THE TITLES OF JOURNALS

(One-word titles are never abbreviated. The word "the" is not used, nor its equivalent in any other language. The word "of" or its equivalent in other languages is used only to discriminate what would otherwise be identical titles in different languages. The word "and" is always used, but indicated by "&" in the Roman alphabet. Only English words are indicated here, but the corresponding words in other languages should receive a corresponding abbreviation. All abbreviations and all one-word titles should be in italics.)

Abnormal .....	<i>Abn.</i>	Japanese .....	<i>Jap.</i>
Abstracts .....	<i>Abst.</i>	Journal .....	<i>J.</i>
American .....	<i>Amer.</i>	Mathematical .....	<i>Math.</i>
Anatomy .....	<i>Anat.</i>	Measurement .....	<i>Meas.</i>
Animal .....	<i>Anim.</i>	Medical .....	<i>Med.</i>
Applied .....	<i>Appl.</i>	Mental .....	<i>Ment.</i>
Archives .....	<i>Arch.</i>	Monographs .....	<i>Monog.</i>
Association .....	<i>Assoc.</i>	Neurology .....	<i>Neurol.</i>
Attitude .....	<i>Attit.</i>	Opinion .....	<i>Opin.</i>
Australian .....	<i>Aust.</i>	Orthopsychiatry .....	<i>Orthopsychiat.</i>
Behavior .....	<i>Behav.</i>	Personality .....	<i>Personal.</i>
British .....	<i>Brit.</i>	Personnel .....	<i>Person.</i>
Bulletin .....	<i>Bull.</i>	Philosophy .....	<i>Philos.</i>
Bureau .....	<i>Bur.</i>	Physics .....	<i>Phys.</i>
Canadian .....	<i>Can.</i>	Physiology .....	<i>Physiol.</i>
Character .....	<i>Charac.</i>	Proceedings .....	<i>Proc.</i>
Children .....	<i>Child.</i>	Psychiatry .....	<i>Psychiat.</i>
Chinese .....	<i>Chin.</i>	Psychoanalysis .....	<i>Psychoanal.</i>
Clinical .....	<i>Clin.</i>	Psychology .....	<i>Psychol.</i>
College .....	<i>Coll.</i>	Psychosomatic .....	<i>Psychosomat.</i>
Comparative .....	<i>Comp.</i>	Quarterly .....	<i>Quart.</i>
Consulting .....	<i>Consult.</i>	Religious .....	<i>Relig.</i>
Contributions .....	<i>Contrib.</i>	Research .....	<i>Res.</i>
Development .....	<i>Devel.</i>	Review .....	<i>Rev.</i>
Educational .....	<i>Educ.</i>	School .....	<i>Sch.</i>
Experimental .....	<i>Exper.</i>	Science .....	<i>Sci.</i>
General .....	<i>Gen.</i>	Social .....	<i>Soc.</i>
Genetic .....	<i>Genet.</i>	Statistics .....	<i>Stat.</i>
Indian .....	<i>Ind.</i>	Studies .....	<i>Stud.</i>
Industrial .....	<i>Indus.</i>	Teacher .....	<i>Teach.</i>
International .....	<i>Internat.</i>	University .....	<i>Univ.</i>
Italian .....	<i>Ital.</i>		

# Preparation of Manuscripts for The Journal Press

## GENERAL INSTRUCTIONS

1. The proper sequence for the parts of your submitted manuscript is as follows: (a) text, (b) references, (c) footnotes, (d) tables, (e) figures, and (f) figure legends. However, monographs start with a table of contents and may have an acknowledgment page before the text and an appendix immediately after the text.
2. Use heavy typewriter paper,  $8\frac{1}{2} \times 11$  inches, double-space *all* lines, and leave margins for editorial work. Do not use onionskin, odd sizes, and abrasive or wax finishes.
3. Submit original typewritten version and one copy. Retain second copy for proofing.
4. Retype any page on which written corrections have been made.
5. Do not begin a sentence with a numeral.
6. A summary at the beginning of the text is required for articles over 500 words.
7. Each quotation should indicate the page number of the original source. The original publisher must give permission for lengthy quotations and use of tables or figures.
8. Do not fold your manuscript.
9. Enclose a submission letter, with a statement that the manuscript is not under consideration elsewhere. If you are unknown to the Editors, kindly give your credentials.

## FORMAT AND SPECIFIC INSTRUCTIONS

### A. TEXT DIVISIONS

#### I. THE TITLES OF JOURNAL ARTICLES AND THE MAJOR SUBDIVISIONS OF MONOGRAPHS ARE PRINTED IN TEN-POINT CAPS CENTERED ON THE PAGE

##### A. THE NEXT SUBDIVISION TITLE IS PRINTED IN CAPS AND SMALL CAPS CENTERED ON THE PAGE

1. *Then Italics, with Principal Words, Upper and Lower Case, Centered on the Page*
- a. *Then italics, upper and lower case, 1-em run-in side head.*
- (1). *Then italics, upper and lower case, 2-em run-in side head.*
- (a). *Then italics, upper and lower case, 3-em run-in side head.*

[Further subdivision should be merged into the text without marginal indentation, and should be numbered with small letters.]

### B. REFERENCES

References should be arranged in alphabetical order by author, numbered and referred to in the text by number (2). Double-space!

The proper form of a book reference is as follows:

2. Doe, J. The Preparation of Manuscripts. New York: Holt, 1963. Pp. 400-418.

The proper form of a journal reference is as follows:

2. Doe, J. The preparation of manuscripts. *J. Gen. Psychol.*, 1963, 68, 450-462.

If in the text it is desirable to refer to a page, thus (2, p. 45).

### C. FOOTNOTES

Use as few as possible and number consecutively in the text thus.<sup>1</sup>

- <sup>1</sup> Footnote (on the separate footnote page). Double-space!

### D. TABLES

Each table should be typed on a separate sheet and should be Arabic numbered (Table 2). Each column requires a heading. Vertical lines should be avoided. Mention tables consecutively in the text and indicate approximate insertion points.

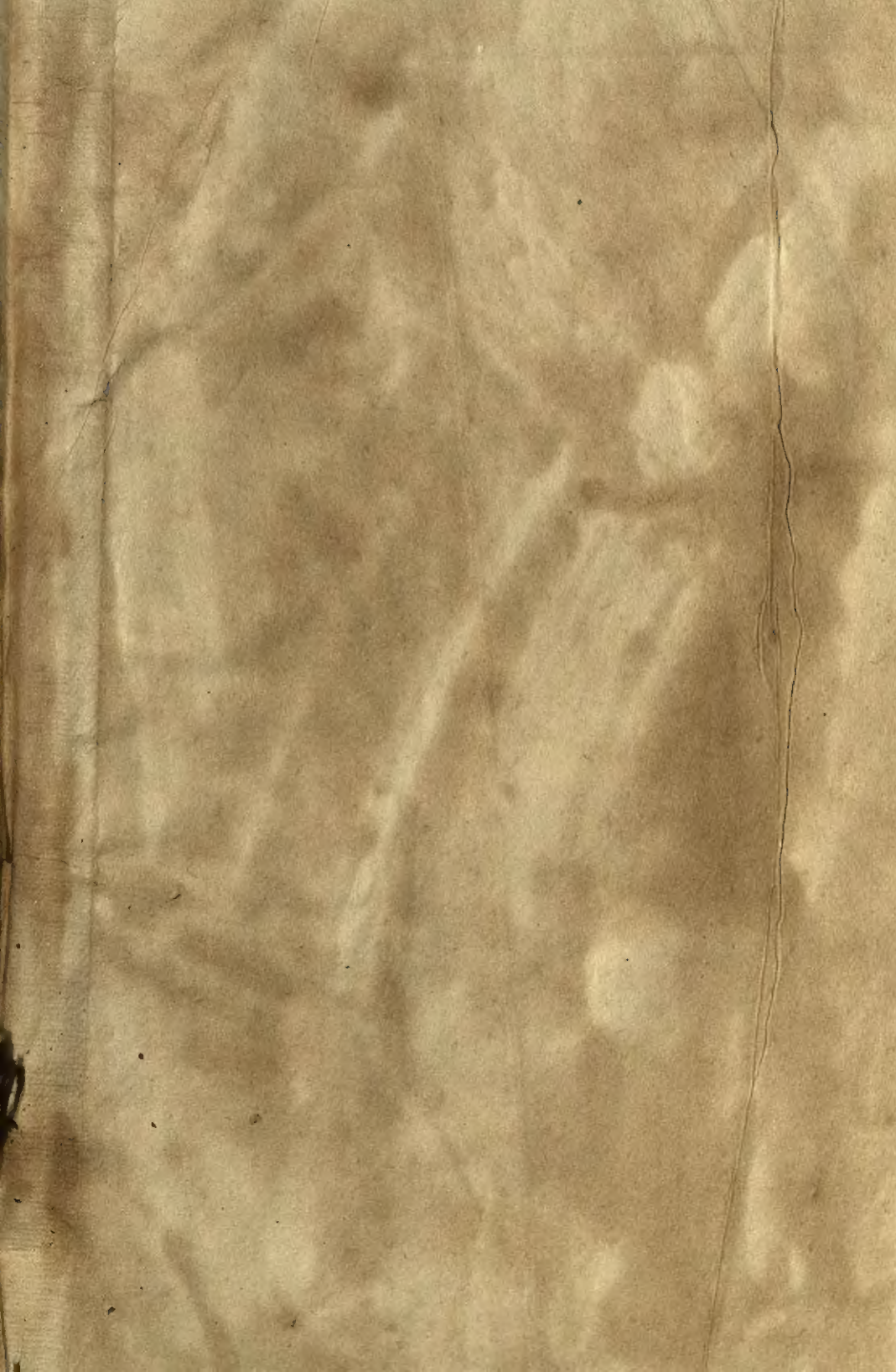
### E. FIGURES

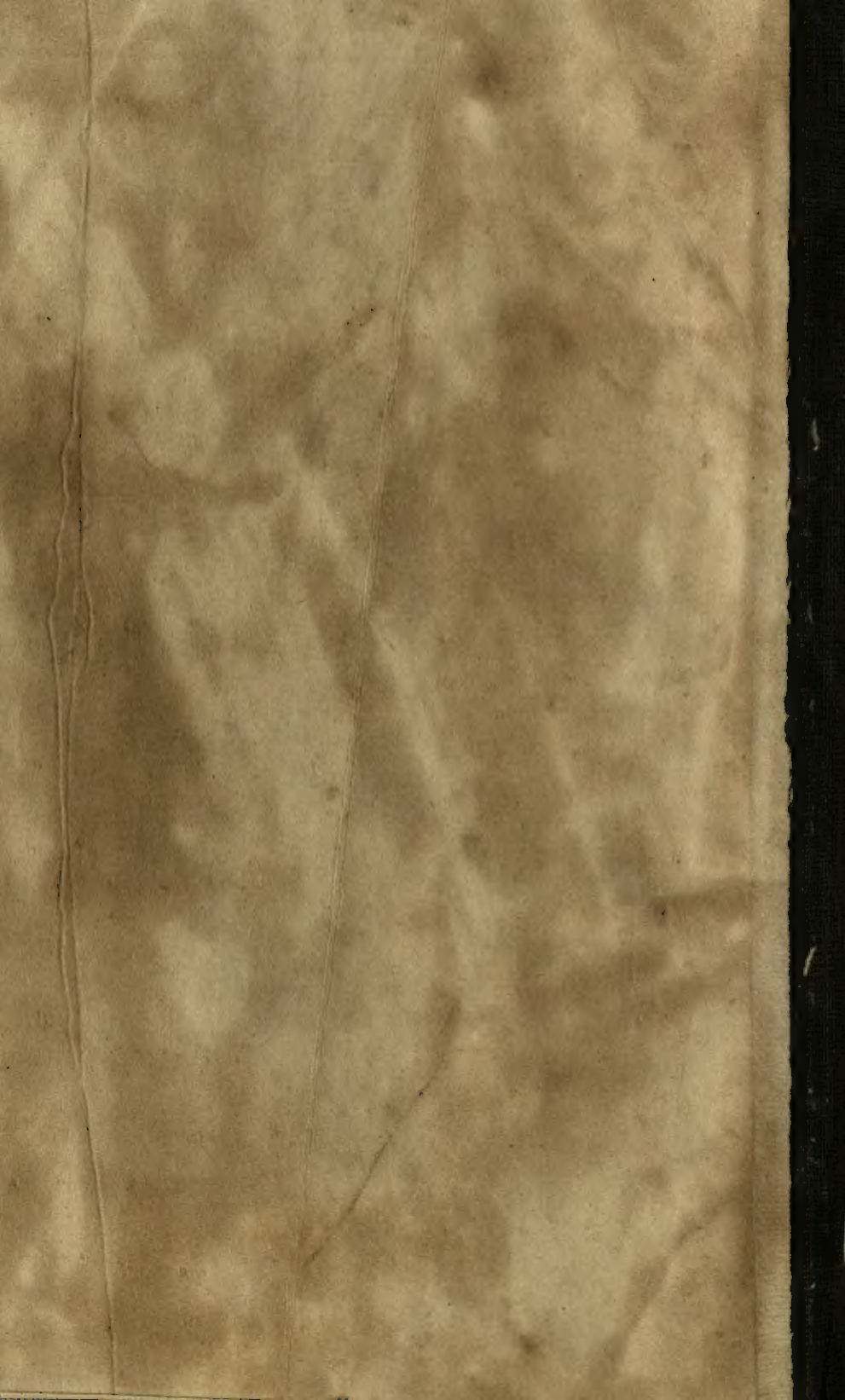
Figures should be submitted as glossy prints of the approximate size for final reproduction (ordinarily  $4-4\frac{1}{2}$  inches in width). Lettering and lines should be sharp and clean. Figures should be Arabic numbered (Figure 2) in the text, and have consecutive mention and approximate insertion points. Each figure requires a legend, but all legends should be submitted double-spaced on a separate Figure Legends page.













MADE BY THE DISABLED  
REHABILITATION - INDIA

CALCUTTA-700 053

